# On the fulfillment of the complementary approximate Karush-Kuhn-Tucker conditions and algorithmic applications

Renan W. Prado[1], Sandra A. Santos[1], and Lucas E. A. Simões[1]

[1] *Institute of Mathematics, University of Campinas, Rua Sergio Buarque de Holanda, 651, 13083-859, Campinas, São Paulo, Brazil*

Emails: `renanwilliamprado@gmail.com`, `sandra@ime.unicamp.br`, `simoes.lea@gmail.com`

January 19, 2023

**Abstract**

Focusing on smooth constrained optimization problems, and inspired by the *complementary approximate Karush-Kuhn-Tucker* (CAKKT) conditions, this work introduces the *weighted complementary approximate Karush-Kuhn-Tucker* (WCAKKT) conditions. They are shown to be verified by limit points generated not only by safeguarded augmented Lagrangian methods, but also by inexact restoration methods, inverse and logarithmic barrier methods, and a penalized algorithm for constrained nonsmooth optimization. Under the analyticity of the feasible set description, and resting upon a desingularization result, the new conditions are proved to be equivalent to the CAKKT conditions. The WCAKKT conditions capture the algebraic elements of the desingularization result needed to characterize CAKKT sequences using a weighted complementarity condition that asymptotically sums zero. Due to its generality and strength, the new condition may help to enlighten the practical performance of algorithms in generating CAKKT sequences.

**Keywords:** Nonlinear programming, Sequential optimality conditions, Mathematical programming methods

**AMS Classification:** 90C46, 90C30, 65K05

# 1 Introduction

The present work aims to provide a unified technique for which, under weak conditions, algorithms applied to constrained optimization problems generate limit points satisfying a powerful optimality condition from the literature – the complementary approximate Karush-Kuhn-Tucker (CAKKT) conditions [13].

Whenever a new optimization algorithm is developed to solve a certain class of problems, the associated convergence analysis must come along with it. In this sense, one usually provides reasonable conditions ensuring the Karush-Kuhn-Tucker (KKT) conditions at limit points [1, 5, 6, 7, 8, 26, 29, 30]. In general, a first-order optimality condition is a proposition of the form: “KKT or not-CQ”, in which CQ is a constraint qualification that involves first-order derivatives of the constraints. Consequently, the strength of an optimality condition is associated with the weakness of the associated CQ. On the other hand, the practicality of an optimality condition may be related to the existence of algorithms whose limit points satisfy it. From this perspective, for problems in which the KKT conditions are not valid at local minimizers, it is valuable to investigate the behavior of the limit points generated by the algorithm under analysis.

To address such an issue, the so-called *sequential optimality conditions* (SOCs) have recently received attention. The first SOC devised was the *approximate gradient projection* (AGP) condition [40]. One of the best properties of SOCs is that, unlike the KKT conditions, they are satisfied by all minimizers regardless of the algebraic description of the feasible set. Therefore, SOCs are considered genuine necessary optimality conditions [1, 8, 13], and they can be seen as a generalization of the KKT conditions. Besides, SOCs are related to stopping criteria for nonlinear programming algorithms.

In addition, SOCs ensure the KKT conditions at limit points under conditions weaker than the Mangasarian-Fromovitz constraint qualification (MFCQ), see [8, Fig. 4.1]. Hence, proving the global convergence for an algorithm under weak constraint qualifications can be reduced to verifying that its limit points satisfy a given SOC. In other words, SOCs are an excellent way of unifying theoretical convergence results of algorithms.

This work focuses on the fulfillment of the CAKKT conditions. Regarding their strength, the conditions are sufficient for optimality in convex problems [13]. Further, for non-convex constrained problems, the fulfillment of the KKT conditions is obtained under the CAKKT-regularity, one of the weakest known constraint qualifications (see [11] for a comparison with the best-known ones). Since, generally, SOCs are not sufficient to guarantee local optimality, the stronger the sequential condition, the better. Bearing this in mind, the CAKKT conditions are stronger than the well-known *approximate Karush-Kuhn-Tucker conditions* (AKKT) [4], thus avoiding undesirable candidates for local minimizers that might be acceptable by AKKT. Hence, in the search for first-order stationarity, CAKKT points have more chance to be local minimizers than AKKT points.

For practical methods, the CAKKT conditions have been put to the proof several times. They were used to ensure strong convergence results of algorithms like safeguarded augmented Lagrangian methods [13]; a penalized algorithm for constrained nonsmooth optimization (PACNO) with an additional necessary condition for local minimizers [43]; a primal-dual augmented Lagrangian method [14]; an augmented Lagrangian algorithm for second-order cone programming and nonlinear semidefinite programming [3], and, recently, an inexact restoration method [15]. For most of the aforementioned algorithms, the CAKKT conditions have shown a favorable convergence theory matching the experimental results. In the current work, the condition will be proven to hold for additional methods.

Moreover, the CAKKT conditions have been widely used to attest optimality for difficult optimization problems with specific structure, such as mathematical programs with complementarity constraints [9]; nonlinear symmetric cone programming and second-order cone programming [2]; nonlinear semidefinite programming [3]; multiobjective optimization problems [41], and cardinality-constrained optimization problems [34].

When it comes to the new optimality condition devised, namely the *Weighted Complementary Approximate KKT* (WCAKKT), the contributions are the relationships with the CAKKT conditions, the applicability to additional algorithms, and further consequences. Up to now, the only way to ensure the validity of the CAKKT conditions for a method would be directly verifying them for the particular method. In this work, an easier alternative has been devised. As briefly discussed in Section 2, some stopping criteria for optimization algorithms are associated with results stronger than others, meaning that the stopping criteria may guarantee that some algorithms have better chances of finding local minimizes than others. Like most of the SOCs, the CAKKT conditions are related to stopping criteria. The WCAKKT conditions come into play as a theoretical assurance that any method that verifies them can be terminated by the stopping criteria related to the CAKKT conditions. This theoretical result also guarantees that both conditions turn out to be equivalent under suitable conditions.

The practical contribution of the new condition is established by proving that different methods satisfy the CAKKT conditions. Among them are a safeguarded augmented Lagrangian [13]; an inexact restoration method [29]; an inverse barrier method [17]; a logarithmic barrier method [12], and the penalized algorithm for constrained nonsmooth optimization (PACNO) applied to smooth problems [32, 33]. Thus, all these methods can be terminated by stopping criteria related to the CAKKT conditions, which is a new result. Moreover, as WCAKKT conditions are equivalent to the CAKKT conditions, we have provided a unified way of capturing the necessary hypothesis for proving the CAKKT conditions, without making use of the Kurdyka-Lojasiewicz inequality, or any other specific relationship associated with the method. In our view, this is a fair compromise to achieve results for distinct methods in our analysis. Three new and noteworthy corollaries should be pointed out. First, the new condition provides a tool ensuring that whenever a nonlinear programming algorithm applied to any instance of a broad class of convex optimization problems generates a convergent sequence satisfying the WCAKKT conditions, the limit point is a solution for the problem – in general, this result cannot be easily achieved just through the CAKKT conditions. Second, it serves to prove that, in the inequality-constrained case, the same sequence that verifies the AGP property [40] also assures the CAKKT conditions, which was not proved before. Third, it guarantees that methods fulfilling the WCAKKT conditions also ensure the KKT condition under CAKKT-regularity [11].

The roadmap of this work is the following. Section 2 contains some of the basic terminologies, with the discussion on how stopping criteria are related to SOCs. In Section 3, the WCAKKT sequential optimality conditions are defined, and the main result of this article is precisely stated. Section 4 is dedicated to the proof of the main theorem, with a discussion about the tightness of the new result. In Section 5, the WCAKKT conditions are related to several mathematical programming algorithms. Our final remarks are stated in Section 6.

**Notation**

- For $\boldsymbol{v} \in \mathbb{R}^n$, $\boldsymbol{v}_+ = (\max\{0, v_1\}, \dots, \max\{0, v_n\})$.
- For $\boldsymbol{v} \in \mathbb{R}^n$, $\boldsymbol{v}_- = (\min\{0, v_1\}, \dots, \min\{0, v_n\})$.
- $\|\cdot\|$ denotes the Euclidean norm.
- $\mathcal{B}[\boldsymbol{x}, \epsilon]$ is the closed ball with center $\boldsymbol{x}$, radius $\epsilon$, and the Euclidean distance.
- $\mathbb{R}_+$ stands for the set of non-negative real numbers.
- $\mathbb{R}_+^*$ denotes the set of strictly positive real numbers.
- The set of natural numbers is denoted by $\mathbb{N}$.
- Given a natural $\ell$, $I_\ell = \{k \in \mathbb{N} : 1 \leq k \leq \ell\}$.
- $\mathbb{N} = \{0, 1, 2, \dots\}$ and $\mathbb{N}^* = \{1, 2, \dots\}$.
- Given any two subsets $A$ and $B$ of $\mathbb{N}$, the notation $A \subset_\infty B$ means that $A$ is an infinite subset of $B$.
- Given $\mathcal{K} \subset_\infty \mathbb{N}$, the notation $\{\boldsymbol{w}^k\}_{k\in\mathcal{K}}$ is used to denote the subsequence of $\{\boldsymbol{w}^k\}_{k\in\mathbb{N}}$ with indexes $k \in \mathcal{K}$, so that the terms indexed within the set $\mathbb{N} \setminus \mathcal{K}$ are ignored. Moreover, $\lim\limits_{k\in\mathcal{K}, k\to\infty}$ will be denoted by $\lim\limits_{k\in\mathcal{K}}$.

# 2 Preliminaries

This study addresses the following optimization problem

$$\begin{aligned} &\min_{\boldsymbol{x}\in\mathbb{R}^n} f(\boldsymbol{x}) \\ &\text{s.t. } \boldsymbol{c}(\boldsymbol{x}) \leq \boldsymbol{0}, \quad \boldsymbol{h}(\boldsymbol{x}) = \boldsymbol{0}, \end{aligned} \tag{1}$$

where $f : \mathbb{R}^n \to \mathbb{R}$, $\boldsymbol{c} : \mathbb{R}^n \to \mathbb{R}^p$ and $\boldsymbol{h} : \mathbb{R}^n \to \mathbb{R}^q$ are all differentiable functions. For simplicity, the set $\Omega := \{\boldsymbol{x} \in \mathbb{R}^n : \boldsymbol{c}(\boldsymbol{x}) \leq \boldsymbol{0} \text{ and } \boldsymbol{h}(\boldsymbol{x}) = \boldsymbol{0}\}$ will stand for the feasible set of problem (1) throughout the entire text.

A mathematical property $\mathcal{P}$ is said to be an *optimality condition* for problems with the above formulation if, given any local minimizer $\boldsymbol{x}^*$ of (1), the property $\mathcal{P}$ holds for $\boldsymbol{x}^*$. For example, "$\boldsymbol{x}^* \in \Omega$" is an optimality condition, but it is very weak, and thus, not useful. Regarding optimality conditions, the stronger the better.

Among the optimality conditions with practical applications, two types can be highlighted: the sequential optimality conditions and the pointwise ones. The pointwise optimality conditions rely on a single feasible point of the given optimization problem, whereas sequential optimality conditions rely on a sequence of – not necessarily feasible – points. A well-known example of a pointwise optimality condition is given by the Fritz John conditions [17, §3.3.5]. For completeness, a point $\boldsymbol{x}^* \in \mathbb{R}^n$ fulfills the Fritz John conditions for problem (1) whenever there exist $\bar{\mu}_0 \in \mathbb{R}$, $\bar{\boldsymbol{\mu}} \in \mathbb{R}^p$, and $\bar{\boldsymbol{\lambda}} \in \mathbb{R}^q$ such that

$$\begin{aligned} \bar{\mu}_0 \nabla f(\boldsymbol{x}^*) + \textstyle\sum_{i=1}^{p} \bar{\mu}_i \nabla c_i(\boldsymbol{x}^*) + \sum_{j=1}^{q} \bar{\lambda}_j \nabla h_j(\boldsymbol{x}^*) = \boldsymbol{0}, \\ (\bar{\mu}_0, \bar{\boldsymbol{\mu}}) \geq (0, \boldsymbol{0}), (\bar{\mu}_0, \bar{\boldsymbol{\mu}}, \bar{\boldsymbol{\lambda}}) \neq (0, \boldsymbol{0}, \boldsymbol{0}), \\ \boldsymbol{c}(\boldsymbol{x}^*) \leq \boldsymbol{0}, \quad \boldsymbol{h}(\boldsymbol{x}^*) = \boldsymbol{0}, \\ \min\{-c_i(\boldsymbol{x}^*), \bar{\mu}_i\} = 0, \text{ for all } i \in I_p. \end{aligned}$$

Under the Mangasarian-Fromovitz Constraint Qualification (MFCQ), the above set of conditions reduces to the well-known KKT conditions. Unfortunately, in 'highly degenerated problems', like the one given in equation (1.2) of [14], the Fritz John conditions hold for infinitely many nonoptimal feasible points, thus being disqualified as an optimality test.

Since optimality conditions measure the adequacy of a point in being a solution, it is desirable to ensure them to be valid at limit points of algorithms. Like the KKT conditions, most pointwise conditions can only be classified as optimality conditions under some type of constraint qualification. This fact reduces the class of problems for which pointwise conditions can be applied since many important optimization problems are intrinsically degenerated: mathematical programs with complementarity constraints [9] and cardinality-constrained problems [37] are examples of such cases. On the other hand, sequential optimality conditions do not need any kind of constraint qualification to be valid at local minimizers and, consequently, they can be applied to a much broader class of optimization problems [9].

Another advantage of sequential optimality conditions over the pointwise ones is their role as natural stopping criteria of optimization algorithms. Indeed, optimization methods are not usually guaranteed to generate limit points satisfying the KKT conditions, but only approximations. That is, for most algorithms, it must be checked whether the current candidate for a minimizer, say $\boldsymbol{x} \in \mathbb{R}^n$, together with the associated Lagrange multipliers, say $\boldsymbol{\mu} \in \mathbb{R}^p_+$ and $\boldsymbol{\lambda} \in \mathbb{R}^q$, ensures that the KKT measurement error reaches a tolerance threshold

$\epsilon > 0$:

$$\left\| \nabla f(\boldsymbol{x}) + \textstyle\sum_{i=1}^{p} \mu_i \nabla c_i(\boldsymbol{x}) + \sum_{j=1}^{q} \lambda_j \nabla h_j(\boldsymbol{x}) \right\| \le \epsilon, \tag{2}$$

$$\|\boldsymbol{c}(\boldsymbol{x})_+\| \le \epsilon, \quad \|\boldsymbol{h}(\boldsymbol{x})\| \le \epsilon, \tag{3}$$

$$\text{and, for all } i \in I_p, |\min\{-c_i(\boldsymbol{x}), \mu_i\}| \le \epsilon. \tag{4}$$

For an algorithm generating a convergent sequence, this stopping criterion is based on the expectancy that the error measured in the KKT conditions, expressed in (2)–(4), converges to zero – with the approximate Lagrange multipliers possibly going to infinity. If such expectancy of convergence is fulfilled, the SOCs derived are called *approximate Karush-Kuhn-Tucker* (AKKT) conditions [4], defined next for the reader's convenience.

**Definition 2.1** (Approximate Karush-Kuhn-Tucker conditions (AKKT) [4])**.** *Consider problem* (1) *and let* $\boldsymbol{x}^*$ *be one of its feasible points. We say that* $\boldsymbol{x}^*$ *is an Approximate KKT (AKKT) point, or satisfies the AKKT conditions, if, for every* $\delta > 0$, $\epsilon > 0$, *there exist* $\boldsymbol{x} \in \mathcal{B}[\boldsymbol{x}^*, \delta] \subset \mathbb{R}^n$ *and approximate Lagrange multipliers* $\boldsymbol{\mu} \in \mathbb{R}^p_+$ *and* $\boldsymbol{\lambda} \in \mathbb{R}^q$ *satisfying* (2)*–*(4)*.*

*Equivalently, in a sequential formulation, the feasible* $\boldsymbol{x}^*$ *is an AKKT point whenever there exist sequences* $\{\boldsymbol{x}^k\}_{k\in\mathbb{N}} \subset \mathbb{R}^n$, $\{\boldsymbol{\mu}^k\}_{k\in\mathbb{N}} \subset \mathbb{R}^p_+$ *and* $\{\boldsymbol{\lambda}^k\}_{k\in\mathbb{N}} \subset \mathbb{R}^q$ *such that* $\boldsymbol{x}^k \to \boldsymbol{x}^*$*,*

$$\lim_{k\to\infty} \left( \nabla f(\boldsymbol{x}^k) + \textstyle\sum_{i=1}^{p} \mu_i^k \nabla c_i(\boldsymbol{x}^k) + \sum_{j=1}^{q} \lambda_j^k \nabla h_j(\boldsymbol{x}^k) \right) = \boldsymbol{0}, \tag{5}$$

$$\textit{and, for all } i \in I_p, \lim_{k\to\infty} \min\{-c_i(\boldsymbol{x}^k), \mu_i^k\} = 0. \tag{6}$$

*In such a context, the primal sequence* $\{\boldsymbol{x}^k\}_{k\in\mathbb{N}}$ *is called an AKKT sequence, and the sequences* $\{\boldsymbol{\mu}^k\}_{k\in\mathbb{N}}$ *and* $\{\boldsymbol{\lambda}^k\}_{k\in\mathbb{N}}$ *are called AKKT multipliers.*

An optimization method is said to be *compatible* with the AKKT conditions (cf. [12]) if it generates a subsequence of iterates and associated approximate Lagrange multipliers satisfying the AKKT conditions either are generated or may be computed – for instance, the interior-point method presented in [12] is an example of such a method. It is important to notice that the aforementioned notion of compatibility does not involve the entire sequence generated by the method, as many algorithms might not converge to a single limit point. Methods compatible with the AKKT conditions inherit the global convergence result of the AKKT conditions. This means, for example, that any algorithm compatible with AKKT conditions converges to KKT points under the weakest constraint qualification associated with AKKT, namely the cone-continuity property (CCP) (see [10] or AKKT-regularity, cf. [11]).

The strength of a sequential optimality condition is associated with its ability to avoid stationary points that are not local minimizers. Despite the practical and intuitive appeal of the AKKT conditions, they may not be appropriate for some intricate problems – see, for example, the simple MPCC instance (3) in [1, Introduction]. Fortunately, this can be avoided by strengthening the AKKT conditions with the following extra requirements:

$$|\mu_i c_i(\boldsymbol{x})| \leq \epsilon, \forall i \in I_p \quad \text{and} \quad |\lambda_j h_j(\boldsymbol{x})| \leq \epsilon, \forall j \in I_q, \tag{7}$$

in which $\epsilon > 0$ is as in (2)–(4). The SOC that emerges from the addition of the above conditions is known as *Complementary AKKT* (CAKKT) [13]. It is stronger than AKKT and tends to find better candidates for minimizers. Its definition is presented next for prompt reference.

**Definition 2.2** (Complementary approximate Karush-Kuhn-Tucker (CAKKT) conditions)**.** *Consider problem* (1) *and let* $\boldsymbol{x}^*$ *be one of its feasible points. The point* $\boldsymbol{x}^*$ *is said a CAKKT point whenever there exist sequences* $\{\boldsymbol{x}^k\}_{k\in\mathbb{N}} \subset \mathbb{R}^n$*,* $\{\boldsymbol{\mu}^k\}_{k\in\mathbb{N}} \subset \mathbb{R}^p_+$ *and* $\{\boldsymbol{\lambda}^k\}_{k\in\mathbb{N}} \subset \mathbb{R}^q$ *such that* $\boldsymbol{x}^k \to \boldsymbol{x}^*$*, and* (5) *holds, together with*

$$\lim_{k\to\infty} \mu_i^k c_i(\boldsymbol{x}^k) = 0, \text{ for all } i \in I_p, \text{ and } \lim_{k\to\infty} \lambda_j^k h_j(\boldsymbol{x}^k) = 0, \text{ for all } j \in I_q. \tag{8}$$

*The resultant sequence* $\{\boldsymbol{x}^k\}_{k\in\mathbb{N}}$ *is called a CAKKT sequence and its associated multipliers* $\{\boldsymbol{\lambda}^k\}_{k\in\mathbb{N}}$ *and* $\{\boldsymbol{\mu}^k\}_{k\in\mathbb{N}}$ *are the CAKKT multipliers.*

Similar to the AKKT conditions, algorithms that generate subsequences of iterates for which there exist associated approximate Lagrange multipliers –either produced by the algorithm or computed afterwards– satisfying the CAKKT conditions are said to be *compatible with the CAKKT conditions*. For such algorithms, the stopping criterion established by the expressions (2)–(4) and (7) is eventually verified. In this study, we are particularly interested in methods that are compatible with CAKKT conditions.

# 3 The new sequential optimality condition and the main result

Sequential optimality conditions, by its very definition, must be satisfied at all local minimizers of (1). Despite this, optimization methods might not be compatible with some sequential optimality conditions, i.e., the generated subsequence of iterates, together with the associated approximate Lagrange multipliers, might not fulfill all the properties of the SOC at hand. Therefore, identifying the algorithms that are compatible with a particular SOC is crucial to understanding the applicability of such a SOC. When it comes to the study of the CAKKT conditions, it is generally related to safeguarded augmented Lagrangian methods [13, 14, 31], and, to the best of our knowledge, there is no unified way of ensuring that CAKKT points are generated by algorithms that do not belong to the class of augmented Lagrangian methods. The present study is devoted to broadening the set of optimization methods that are compatible with the CAKKT conditions.

It is well known that both safeguarded augmented Lagrangian methods and inexact restoration methods generate limit points satisfying the sequential optimality condition known as approximate gradient projection (AGP) [40].

However, only safeguarded augmented Lagrangian methods are guaranteed to generate sequences verifying the CAKKT conditions. Since CAKKT implies AGP [13], identifying the connections between these two conditions may give us a hint of what is needed for a method to be compatible with CAKKT. As it will be seen later, the connection between AGP and CAKKT can be made by a new sequential optimality condition called *Weighted Complementary Approximate Karush-Kuhn-Tucker (WCAKKT)*, defined as follows.

**Definition 3.1** (Weighted complementary approximate KKT (WCAKKT))**.** *Consider problem* (1) *and let* $\boldsymbol{x}^*$ *be an AKKT point associated with the AKKT primal sequence* $\{\boldsymbol{x}^k\}_{k\in\mathbb{N}} \subset \mathbb{R}^n$ *and AKKT multipliers* $\{\boldsymbol{\mu}^k\}_{k\in\mathbb{N}} \subset \mathbb{R}^p_+$ *and* $\{\boldsymbol{\lambda}^k\}_{k\in\mathbb{N}} \subset \mathbb{R}^q$*. We say that the point* $\boldsymbol{x}^*$ *satisfies the* weighted CAKKT (WCAKKT) conditions *if, additionally, the following implication holds:*

**(W)** for all $\mathcal{K} \subset_\infty \mathbb{N}$, *and* for all *sequences* $\{\boldsymbol{\nu}^k\}_{k\in\mathcal{K}} \subset \mathbb{R}^p$ *and* $\{\boldsymbol{\eta}^k\}_{k\in\mathcal{K}} \subset \mathbb{R}^q$ *satisfying* $\lim_{k\in\mathcal{K}} \boldsymbol{\nu}^k = \boldsymbol{\nu} \in (\mathbb{N}^*)^p$, $\lim_{k\in\mathcal{K}} \boldsymbol{\eta}^k = \boldsymbol{\eta} \in (\mathbb{N}^*)^q$ *and*

$$\lim_{k\in\mathcal{K}} \left(\sum_{i=1}^p \nu_i^k \mu_i^k c_i(\boldsymbol{x}^k) + \sum_{j=1}^q \eta_j^k \lambda_j^k h_j(\boldsymbol{x}^k)\right) = 0, \text{ necessarily}$$

$$\lim_{k\in\mathcal{K}} \mu_i^k c_i(\boldsymbol{x}^k) = 0, \text{ for all } i \in I_p, \text{ and } \lim_{k\in\mathcal{K}} \lambda_j^k h_j(\boldsymbol{x}^k) = 0, \text{ for all } j \in I_q.$$

*The resultant sequence* $\{\boldsymbol{x}^k\}_{k\in\mathbb{N}}$ *is called WCAKKT sequence, and* $\{\boldsymbol{\mu}^k\}_{k\in\mathbb{N}}$ *and* $\{\boldsymbol{\lambda}^k\}_{k\in\mathbb{N}}$ *are called WCAKKT multipliers. An algorithm is said to be compatible with the WCAKKT conditions whenever it generates a primal sequence, and sequences of approximate Lagrange multipliers either are generated or can be obtained such that, at least for a subsequence, the former can be regarded as a WCAKKT sequence, and the latter, as the associated WCAKKT multipliers.*

**Remark 3.1.** *It is worth mentioning that the sequences* $\{\boldsymbol{\nu}^k\}$ *and* $\{\boldsymbol{\eta}^k\}$ *provide flexibility to address methods with distinct features. Moreover, constant vectors* $\boldsymbol{\nu} \in (\mathbb{N}^*)^p$ *and* $\boldsymbol{\eta} \in (\mathbb{N}^*)^q$ *are not enough to attain the result of Theorem 4.1 ahead. As a counter-example, let the sequences* $r_1^k := \mu_1^k c_1(\boldsymbol{x}^k) = k - 2$, $r_2^k := \mu_2^k c_2(\boldsymbol{x}^k) = -k$, $\nu_1^k = 1 + 1/r_1^k$, $\nu_2^k = 1 + 1/r_2^k$, *and observe that* $\lim_{k\to\infty} \nu_1^k = \lim_{k\to\infty} \nu_2^k = 1$*. Moreover, for all* $k \in \mathbb{N}$*, notice that* $\nu_1^k r_1^k + \nu_2^k r_2^k = 0$*, but* $r_1^k + r_2^k = \mu_1^k c_1(\boldsymbol{x}^k) + \mu_2^k c_2(\boldsymbol{x})^k = -2$*. Therefore, by replacing* $\nu_1^k$ *and* $\nu_2^k$ *for all* $k$ *by the respective limit values on* $\nu_1^k r_1^k + \nu_2^k r_2^k$*, the weighted sum is no longer zero.*

To identify the link between CAKKT and AGP, we first establish a result ensuring that every CAKKT sequence is a WCAKKT one.

**Proposition 3.1.** *For problem* (1)*, if* $\boldsymbol{x}^*$ *is a CAKKT point, every CAKKT sequence converging to* $\boldsymbol{x}^*$ *is a WCAKKT sequence.*

*Proof.* The result follows trivially from the fact that any sequence satisfying CAKKT fulfills (5) and (8), which verify AKKT and the implication (**W**). □

In the scope of problem (1), it is known that the AGP conditions are stronger than the AKKT conditions [4, Counter-example 3.1.]. Furthermore, it has been already established that AGP is strictly weaker than CAKKT [13]. For problems described just by inequality constraints, however, no relationship between AGP and CAKKT had been established yet. Aiming at addressing this matter, and under the assumption that the inequality constraints are stated by analytic functions, the next result grounds such a connection through the WCAKKT conditions. The result linking WCAKKT and CAKKT is important for our purposes because, in the aforementioned context, one can assure that AGP implies WCAKKT, which, therefore, guarantees that AGP implies CAKKT. The statement of the result is given below, but its proof is deferred to the next section.

**Theorem 3.1** (**Main Theorem**). *Assume that the constraints of problem* (1) *are described by analytic functions, and its objective function has bounded derivatives in a neighborhood of the feasible point* $\boldsymbol{x}^*$. *If* $\boldsymbol{x}^*$ *is a WCAKKT point, then every WCAKKT sequence is a CAKKT sequence, and the associated WCAKKT multipliers are CAKKT multipliers.*

# 4 Analytic functions, desingularization and the proof of the main result

Many results based on sequential optimality conditions assume the validity of the Kurdyka-Lojasiewicz inequality of the infeasibility measure [2, 6, 13, 14]. However, a wide range of optimization problems present properties that ensure stronger conditions than just this general inequality. The literature is plentiful on optimization problems modeling practical instances that have feasible sets described by analytic functions [19, 22, 23, 36, 38]. As a consequence, we have assumed, throughout this section, that the functions describing the constraints in problem (1) are all analytic. This allowed us to link the CAKKT condition not only with an inexact restoration algorithm but also with other well-known optimization methods.

The next result, based on a desingularization strategy [16, 20, 21, 35], builds the foundations for the proof of our main theorem, which will lead to the unification of the convergence analysis of several optimization methods.

**Theorem 4.1.** *Assume that the constraints of the optimization problem* (1) *are described by analytic functions, the point* $\boldsymbol{x}^*$ *is feasible for such a problem and the objective function is locally differentiable around* $\boldsymbol{x}^*$, *with bounded derivatives in such a neighborhood. Additionally, assume that there exist an AKKT sequence* $\{\boldsymbol{x}^k\}_{k\in\mathbb{N}}$ *with corresponding AKKT multipliers* $\{\boldsymbol{\mu}^k\}_{k\in\mathbb{N}}$ *and* $\{\boldsymbol{\lambda}^k\}_{k\in\mathbb{N}}$. *Then, there exist an infinite subset of natural numbers* $\mathcal{K}$ *and sequences of real vectors* $\{\boldsymbol{\nu}^k\}_{k\in\mathcal{K}} \subset \mathbb{R}^p$ *and* $\{\boldsymbol{\eta}^k\}_{k\in\mathcal{K}} \subset \mathbb{R}^q$ *so that* $\lim_{k\in\mathcal{K}} \boldsymbol{\nu}^k = \boldsymbol{\nu} \in (\mathbb{N}^*)^p$, $\lim_{k\in\mathcal{K}} \boldsymbol{\eta}^k = \boldsymbol{\eta} \in (\mathbb{N}^*)^q$ *and*

$$\lim_{k\in\mathcal{K}} \left( \sum_{i=1}^p \nu_i^k \mu_i^k c_i(\boldsymbol{x}^k) + \sum_{j=1}^q \eta_j^k \lambda_j^k h_j(\boldsymbol{x}^k) \right) = 0. \tag{9}$$

*Proof.* This demonstration has been organized in five stages: 1) set up the notation; 2) restate the main relationships to exclude trivial analytic constraints from the analysis; 3) apply [44, Thm. 2.1] to ensure a desingularization of the feasible set description; 4) recall basic properties and definitions of analytic manifolds; 5) proceed with the algebraic manipulations to conclude the proof.

To start with, let $\boldsymbol{x}^*$ be a feasible point of problem (1) that satisfies the AKKT condition. From the Definition 2.1, there exist sequences $\{\boldsymbol{x}^k\}_{k\in\mathbb{N}} \subset \mathbb{R}^n$, $\{\boldsymbol{\mu}^k\}_{k\in\mathbb{N}} \subset \mathbb{R}^p_+$, $\{\boldsymbol{\lambda}^k\}_{k\in\mathbb{N}} \subset \mathbb{R}^q$ and $\{\boldsymbol{v}^k\}_{k\in\mathbb{N}} \subset \mathbb{R}^n$ such that $\boldsymbol{x}^k \to \boldsymbol{x}^*$, $\boldsymbol{v}^k \to 0$,

$$\nabla f(\boldsymbol{x}^k) + \textstyle\sum_{i=1}^{p} \mu_i^k \nabla c_i(\boldsymbol{x}^k) + \sum_{j=1}^{q} \lambda_j^k \nabla h_j(\boldsymbol{x}^k) = \boldsymbol{v}^k, \tag{10}$$

and, if

$$c_i(\boldsymbol{x}^*) < 0, \text{ then } \lim_{k\to\infty} \mu_i^k = 0. \tag{11}$$

Defining the set $\mathcal{I}(\boldsymbol{x}^*) = \{i \in I_p : c_i(\boldsymbol{x}^*) = 0\}$, due to the condition (11), whenever $i \notin \mathcal{I}(\boldsymbol{x}^*)$, i.e., $c_i(\boldsymbol{x}^*) < 0$, then necessarily $\lim_{k\to\infty} \mu_i^k = 0$. In such a case, besides the fulfillment of the complementarity

$$\lim_{k\to\infty} c_i(\boldsymbol{x}^k)\mu_i^k = 0 \tag{12}$$

for $i \notin \mathcal{I}(\boldsymbol{x}^*)$, as $\mu_i^k \to 0$, we may rewrite (10) as follows

$$\begin{aligned} \boldsymbol{w}^k &:= \boldsymbol{v}^k - \textstyle\sum_{i\notin\mathcal{I}(\boldsymbol{x}^*)} \mu_i^k \nabla c_i(\boldsymbol{x}^k) \\ &= \nabla f(\boldsymbol{x}^k) + \textstyle\sum_{i\in\mathcal{I}(\boldsymbol{x}^*)} \mu_i^k \nabla c_i(\boldsymbol{x}^k) + \sum_{j\in I_q} \lambda_j^k \nabla h_j(\boldsymbol{x}^k), \end{aligned} \tag{13}$$

with $\lim_{k\to\infty} \boldsymbol{w}^k = \mathbf{0}$.

In the second part of the proof, it is essential to access the *nontrivial* constraints (i.e., those that do not vanish identically on any connected component of their domains). The following sets refer to neighborhoods $\mathcal{U}_*$ of $\boldsymbol{x}^*$, that are not necessarily the same for distinct constraints, and contain the indices of the trivial constraints:

$$\mathcal{I} := \{i \in \mathcal{I}(\boldsymbol{x}^*) \ : \text{ there exists } \mathcal{U}_* \text{ such that } c_i(\boldsymbol{x}) = 0, \forall \boldsymbol{x} \in \mathcal{U}_*\}, \tag{14}$$

$$\mathcal{J} := \{j \in I_q \ : \text{ there exists } \mathcal{U}_* \text{ such that } h_j(\boldsymbol{x}) = 0, \forall \boldsymbol{x} \in \mathcal{U}_*\}. \tag{15}$$

Since at $\boldsymbol{x}^k$, and for $k$ large enough, the gradients of the constraints indexed by $\mathcal{I}$ and $\mathcal{J}$ are null, it is possible to shorten (13) to the indices that belong to $i \in \mathcal{I}(\boldsymbol{x}^*) \setminus \mathcal{I}$ and $j \in I_q \setminus \mathcal{J}$

$$\nabla f(\boldsymbol{x}^k) + \sum_{i\in\mathcal{I}(\boldsymbol{x}^*)\setminus\mathcal{I}} \mu_i^k \nabla c_i(\boldsymbol{x}^k) + \sum_{j\in I_q\setminus\mathcal{J}} \lambda_j^k \nabla h_j(\boldsymbol{x}^k) = \boldsymbol{w}^k. \tag{16}$$

Moreover, for each $i \in \mathcal{I}$ and $j \in \mathcal{J}$,

$$\mu_i^k c_i(\boldsymbol{x}^k) = 0 \quad \text{ and } \quad \lambda_j^k h_j(\boldsymbol{x}^k) = 0, \tag{17}$$

for large enough $k$, since the constraints involved are identically zero around $\boldsymbol{x}^*$. On the other hand, being described by real analytic functions, none of the constraints $\{c_i\}_{\mathcal{I}(\boldsymbol{x}^*)\setminus\mathcal{I}} \cup \{h_j\}_{j\in I_q\setminus\mathcal{J}}$ may be identically zero in $\mathbb{R}^n$, and thus constitute a set of nontrivial analytic functions. It is worth noticing that the only necessary terms in the weighted complementarity stated in expression (9) are the ones with $i \in \mathcal{I}(\boldsymbol{x}^*) \setminus \mathcal{I}$ and $j \in I_q \setminus \mathcal{J}$ since the remaining terms of the sum vanish, in view of (8). Our effort will be focused on proving that the expression (9) holds with the indexes taken in $\mathcal{I}(\boldsymbol{x}^*) \setminus \mathcal{I}$ and $I_q \setminus \mathcal{J}$ with the proper identification in the summations.

The third part of the proof rests upon Theorem 2.1 of [44] to ensure a desingularization of the set $\{c_i\}_{\mathcal{I}(\boldsymbol{x}^*)\setminus\mathcal{I}} \cup \{h_j\}_{j\in I_q\setminus\mathcal{J}}$. Indeed, being a finite set of nontrivial analytic functions, the aforementioned result assures the existence of an analytic manifold $\mathcal{M}$ of pure dimension $n$ (i.e., such that all the connected components have the same dimension), and a proper (i.e., a continuous function for which the inverse image of compact subsets is compact) and surjective analytic function $\boldsymbol{\nu} : \mathcal{M} \to \mathbb{R}^n$ such that the set $\{c_i \circ \boldsymbol{\nu}\}_{\mathcal{I}(\boldsymbol{x}^*)\setminus\mathcal{I}} \cup \{h_j \circ \boldsymbol{\nu}\}_{j\in I_q\setminus\mathcal{J}}$ is desingularized, in the sense of [44]. Roughly speaking, being desingularized is the same as being able to break the function into a product of an analytic function that never vanishes and a special monomial.

In the fourth stage, we will prepare the grounds for the algebraic manipulations. First, notice that, since the function $\boldsymbol{\nu}$ is surjective, we may set up a sequence $\{\boldsymbol{p}^k\}_{k\in\mathbb{N}} \subset \mathcal{M}$ such that

$$\boldsymbol{\nu}(\boldsymbol{p}^k) = \boldsymbol{x}^k, \quad \text{for all } k \in \mathbb{N}. \tag{18}$$

Now, because the sequence $\{\boldsymbol{x}^k\}_{k\in\mathbb{N}}$ is convergent, and thus bounded, there exists a compact set $\mathcal{C}$ such that $\{\boldsymbol{x}^k\}_{k\in\mathbb{N}} \subset \mathcal{C}$. Consequently, the inverse image $\boldsymbol{\nu}^{-1}(\{\boldsymbol{x}^k\}_{k\in\mathbb{N}})$ is contained in the set $\boldsymbol{\nu}^{-1}(\mathcal{C})$, which is compact, as it is the inverse image of a compact set by the proper function $\boldsymbol{\nu}$. In particular, the sequence $\{\boldsymbol{p}^k\}_{k\in\mathbb{N}}$ is contained in the compact set $\boldsymbol{\nu}^{-1}(\mathcal{C})$. Being the analytical manifold $\mathcal{M}$ a metric space, compact subsets of $\mathcal{M}$ are sequentially compact. As a result, the sequence $\{\boldsymbol{p}^k\}_{k\in\mathbb{N}}$ admits a convergent subsequence

$$\lim_{k\in\mathcal{K}} \boldsymbol{p}^k = \boldsymbol{p}^*. \tag{19}$$

Some notation to be used in the sequel must be defined. The open cube centered at the origin is denoted by

$$C^n(\epsilon) := \{\boldsymbol{x} \in \mathbb{R}^n \,:\, |x_j| < \epsilon,\ j \in I_n\}.$$

Given an integer $\ell$ (which either belongs to $\mathcal{I}(\boldsymbol{x}^*) \setminus \mathcal{I}$ or to $I_q \setminus \mathcal{J}$), an $n$-tuple of nonnegative integers $\boldsymbol{\alpha}^{(\ell)} \in \mathbb{Z}^n_+$, and $\boldsymbol{y} \in C^n(\epsilon)$, the real number

$$\underline{\boldsymbol{y}}^{\boldsymbol{\alpha}^{(\ell)}} := y_1^{\alpha_1^{(\ell)}} y_2^{\alpha_2^{(\ell)}} \cdots y_n^{\alpha_n^{(\ell)}}$$

denotes the product of the powers applied componentwise.

As the set $\{c_i \circ \boldsymbol{\nu}\}_{\mathcal{I}(\boldsymbol{x}^*)\setminus\mathcal{I}} \cup \{h_j \circ \boldsymbol{\nu}\}_{j\in I_q\setminus\mathcal{J}}$ is desingularized, there exist an analytic diffeomorphism $\Xi : C^n(\epsilon) \to \mathcal{V}$, where the set $\mathcal{V} \subset \mathcal{M}$ is open, $\boldsymbol{p}^* \in \mathcal{V}$, $\mathbf{0} \in C^n(\epsilon)$, $\Xi(\mathbf{0}) = \boldsymbol{p}^*$, and analytic functions $A_\ell : C^n(\epsilon) \to \mathbb{R}$, with $\ell \in \mathcal{I}(\boldsymbol{x}^*)\setminus\mathcal{I}$ or $\ell \in I_q \setminus \mathcal{J}$, such that

$$c_i \circ \boldsymbol{\nu} \circ \Xi(\boldsymbol{y}) = A_i(\boldsymbol{y})\underline{\boldsymbol{y}}^{\boldsymbol{\alpha}^{(i)}} \quad \text{and} \tag{20}$$

$$h_j \circ \boldsymbol{\nu} \circ \Xi(\boldsymbol{y}) = A_j(\boldsymbol{y})\underline{\boldsymbol{y}}^{\boldsymbol{\alpha}^{(j)}}, \tag{21}$$

for all $\boldsymbol{y} \in C^n(\epsilon)$, $i \in \mathcal{I}(\boldsymbol{x}^*) \setminus \mathcal{I}$ and $j \in I_q \setminus \mathcal{J}$. It is worth noticing that, since the functions $\boldsymbol{\nu}$ and $\Xi$ are analytic, the composition $\boldsymbol{\nu} \circ \Xi$ is analytic as well. Moreover, since the expressions given in (20) and (21) assume value zero at $\boldsymbol{y} = \mathbf{0}$ (because of the feasibility of $\boldsymbol{x}^*$), it holds

$$\textstyle\sum_{\mathrm{i}=1}^{n} \alpha_{\mathrm{i}}^{(\ell)} \geq 1, \text{ for every } \ell \in \mathcal{I}(\boldsymbol{x}^*) \setminus \mathcal{I} \text{ or } \ell \in I_q \setminus \mathcal{J}. \tag{22}$$

Additionally, the $i$-th partial derivative of the application (20) may be stated by

$$\begin{aligned}\frac{\partial}{\partial y_i}(c_\ell \circ \boldsymbol{\nu} \circ \Xi)(\boldsymbol{y}) &= \left(\frac{\partial}{\partial y_i} A_\ell(\boldsymbol{y})\right) \underline{\boldsymbol{y}}^{\boldsymbol{\alpha}^{(\ell)}} + A_\ell(\boldsymbol{y}) \frac{\partial}{\partial y_i} \underline{\boldsymbol{y}}^{\boldsymbol{\alpha}^{(\ell)}} \\ &= \left(\frac{\partial}{\partial y_i} A_\ell(\boldsymbol{y})\right) \underline{\boldsymbol{y}}^{\boldsymbol{\alpha}^{(\ell)}} + \alpha_i^{(\ell)} A_\ell(\boldsymbol{y}) y_i^{\alpha_i^{(\ell)}-1} \prod_{t\in I_n\setminus\{i\}} y_t^{\alpha_t^{(\ell)}}.\end{aligned} \tag{23}$$

On the other hand, by the chain rule, we obtain

$$\frac{\partial}{\partial y_i}(c_\ell \circ \boldsymbol{\nu} \circ \Xi)(\boldsymbol{y}) = \boldsymbol{e}_i^T J_{\boldsymbol{\nu}\circ\Xi}(\boldsymbol{y})^T \nabla c_i((\boldsymbol{\nu} \circ \Xi)(\boldsymbol{y})), \tag{24}$$

in which $\boldsymbol{e}_i \in \mathbb{R}^n$ is the $i$-th canonical vector. Therefore, multiplying (23) and (24) by $y_i$, and adding the obtained expressions yield

$$\begin{aligned}\boldsymbol{y}^T \nabla(c_\ell \circ \boldsymbol{\nu} \circ \Xi)(\boldsymbol{y}) &= \boldsymbol{y}^T J_{\boldsymbol{\nu}\circ\Xi}(\boldsymbol{y})^T \nabla c_\ell((\boldsymbol{\nu} \circ \Xi)(\boldsymbol{y})) \\ &= A_\ell(\boldsymbol{y})\underline{\boldsymbol{y}}^{\boldsymbol{\alpha}^{(\ell)}} \left(\frac{\boldsymbol{y}^T \nabla A_\ell(\boldsymbol{y})}{A_\ell(\boldsymbol{y})} + \sum_{\mathrm{i}=1}^{n} \alpha_{\mathrm{i}}^{(\ell)}\right),\end{aligned} \tag{25}$$

for every $\boldsymbol{y} \in C^n(\epsilon)$, and any $\ell \in \mathcal{I}(\boldsymbol{x}^*) \setminus \mathcal{I}$.

Proceeding in a similar way with respect to the constraint $h_\ell$, for any $\ell \in I_q \setminus \mathcal{J}$ and every $\boldsymbol{y} \in C^n(\epsilon)$, we obtain

$$\begin{aligned}\boldsymbol{y}^T \nabla(h_\ell \circ \boldsymbol{\nu} \circ \Xi)(\boldsymbol{y}) &= \boldsymbol{y}^T J_{\boldsymbol{\nu}\circ\Xi}(\boldsymbol{y})^T \nabla h_\ell((\boldsymbol{\nu} \circ \Xi)(\boldsymbol{y})) \\ &= A_\ell(\boldsymbol{y})\underline{\boldsymbol{y}}^{\boldsymbol{\alpha}^{(\ell)}} \left(\frac{\boldsymbol{y}^T \nabla A_\ell(\boldsymbol{y})}{A_\ell(\boldsymbol{y})} + \sum_{\mathrm{j}=1}^{n} \alpha_{\mathrm{j}}^{(\ell)}\right).\end{aligned} \tag{26}$$

Being $\Xi$ a diffeomorphism, and since $\boldsymbol{p}^k \in \mathcal{V}$, for $k \in \mathcal{K}$ and sufficiently large, from (19) we may define

$$\boldsymbol{y}^k := \Xi^{-1}(\boldsymbol{p}^k), \tag{27}$$

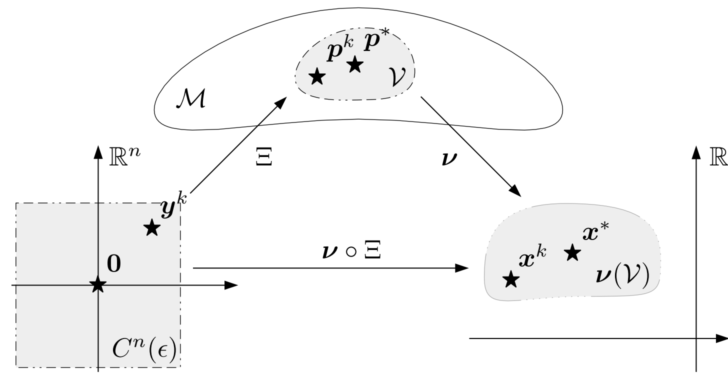


Figure 1: Illustration of the elements related to the sequence $\{\boldsymbol{y}^k\}_{k\in\mathcal{K}}$. Although the sets $C^n(\epsilon)$ and $\mathcal{V}$ are open, the set $\boldsymbol{\nu}(\mathcal{V})$ is not guaranteed to be open, and the lighter colour of its boundary indicates this fact.

with $\boldsymbol{y}^k \in C^n(\epsilon)$, for $k \in \mathcal{K}$ and large enough. The ideas behind such a sequence are depicted in Figure 1. Observe that $\lim_{k\in\mathcal{K}} \boldsymbol{y}^k = \Xi^{-1}(\boldsymbol{p}^*) = \boldsymbol{0}$, as continuity implies sequential continuity in analytic manifolds. In other words, given $\epsilon' > 0$, as the set $C^n(\epsilon')$ is open, then except possibly for a finite set of indices $k$ in $\mathcal{K}$, we have $\boldsymbol{y}^k \in C^n(\epsilon')$. Hence,

$$\lim_{k\in\mathcal{K}} \|\boldsymbol{y}^k\|_2^2 = 0. \tag{28}$$

Furthermore, by the construction of the sequences $\{\boldsymbol{p}^k\}_{k\in\mathbb{N}}$ in (18) and $\{\boldsymbol{y}^k\}_{k\in\mathcal{K}}$ in (27), it follows that

$$\boldsymbol{x}^k = (\boldsymbol{\nu} \circ \Xi)(\boldsymbol{y}^k). \tag{29}$$

We have thus completed the requirements to start the fifth, and central, stage of the proof. Defining

$$\eta_\ell^k := \left( \frac{(\boldsymbol{y}^k)^T \nabla A_\ell(\boldsymbol{y}^k)}{A_\ell(\boldsymbol{y}^k)} + \sum_{\mathfrak{s}=1}^{n} \alpha_{\mathfrak{s}}^{(\ell)} \right), \tag{30}$$

for $k \in \mathcal{K}_1$ and large enough, we have

$$\begin{aligned} &(\boldsymbol{y}^k)^T J_{\boldsymbol{\nu}\circ\Xi}(\boldsymbol{y}^k)^T \left(\boldsymbol{w}^k - \nabla f(\boldsymbol{x}^k)\right) = \\ &\overset{(16)}{=} (\boldsymbol{y}^k)^T J_{\boldsymbol{\nu}\circ\Xi}(\boldsymbol{y}^k)^T \left( \sum_{i\in\mathcal{I}(\boldsymbol{x}^*)\setminus\mathcal{I}} \mu_i^k \nabla c_i(\boldsymbol{x}^k) + \sum_{j\in I_q\setminus\mathcal{J}} \lambda_j^k \nabla h_j(\boldsymbol{x}^k) \right) \\ &\overset{(29)}{=} (\boldsymbol{y}^k)^T J_{\boldsymbol{\nu}\circ\Xi}(\boldsymbol{y}^k)^T \left( \sum_{i\in\mathcal{I}(\boldsymbol{x}^*)\setminus\mathcal{I}} \mu_i^k \nabla c_i((\boldsymbol{\nu}\circ\Xi)(\boldsymbol{y}^k)) \right. \\ &\qquad \left. + \sum_{j\in I_q\setminus\mathcal{J}} \lambda_j^k \nabla h_j((\boldsymbol{\nu}\circ\Xi)(\boldsymbol{y}^k)) \right) \end{aligned}$$

$$
\begin{aligned}
&\overset{(25)}{\underset{(26)}{=}} \sum_{i\in\mathcal{I}(\boldsymbol{x}^*)\setminus\mathcal{I}} \mu_i^k A_i(\boldsymbol{y}^k)(\underline{\boldsymbol{y}^k})^{\boldsymbol{\alpha}^{(i)}} \left( \frac{(\boldsymbol{y}^k)^T \nabla A_i(\boldsymbol{y}^k)}{A_i(\boldsymbol{y}^k)} + \sum_{\mathrm{i}=1}^n \alpha_{\mathrm{i}}^{(i)} \right) \\
&\quad + \sum_{j\in I_q\setminus\mathcal{J}} \lambda_j^k A_j(\boldsymbol{y}^k)(\underline{\boldsymbol{y}^k})^{\boldsymbol{\alpha}^{(j)}} \left( \frac{(\boldsymbol{y}^k)^T \nabla A_j(\boldsymbol{y}^k)}{A_j(\boldsymbol{y}^k)} + \sum_{\mathrm{j}=1}^n \alpha_{\mathrm{j}}^{(j)} \right) \\
&\overset{(30)}{=} \sum_{i\in\mathcal{I}(\boldsymbol{x}^*)\setminus\mathcal{I}} \eta_i^k \mu_i^k A_i(\boldsymbol{y}^k)(\underline{\boldsymbol{y}^k})^{\boldsymbol{\alpha}^{(i)}} + \sum_{j\in I_q\setminus\mathcal{J}} \eta_j^k \lambda_j^k A_j(\boldsymbol{y}^k)(\underline{\boldsymbol{y}^k})^{\boldsymbol{\alpha}^{(j)}} \\
&\overset{(20)}{\underset{(21)}{=}} \sum_{i\in\mathcal{I}(\boldsymbol{x}^*)\setminus\mathcal{I}} \eta_i^k \mu_i^k c_i((\boldsymbol{\nu}\circ\Xi)(\boldsymbol{y}^k)) + \sum_{j\in I_q\setminus\mathcal{J}} \eta_j^k \lambda_j^k h_j((\boldsymbol{\nu}\circ\Xi)(\boldsymbol{y}^k)) \\
&\overset{(29)}{=} \sum_{i\in\mathcal{I}(\boldsymbol{x}^*)\setminus\mathcal{I}} \eta_i^k \mu_i^k c_i(\boldsymbol{x}^k) + \sum_{j\in I_q\setminus\mathcal{J}} \eta_j^k \lambda_j^k h_j(\boldsymbol{x}^k)
\end{aligned}
$$

Consequently, for $k \in \mathcal{K}_1$ and sufficiently large, we have

$$
\begin{aligned}
&\sum_{i\in\mathcal{I}(\boldsymbol{x}^*)\setminus\mathcal{I}} \eta_i^k \mu_i^k c_i(\boldsymbol{x}^k) + \sum_{j\in I_q\setminus\mathcal{J}} \eta_j^k \lambda_j^k h_j(\boldsymbol{x}^k) \\
&= (\boldsymbol{y}^k)^T J_{\boldsymbol{\nu}\circ\Xi}(\boldsymbol{y}^k)^T \left(\boldsymbol{w}^k - \nabla f(\boldsymbol{x}^k)\right) =: \bar{\epsilon}_k \overset{k\in\mathcal{K}}{\longrightarrow} 0
\end{aligned} \tag{31}
$$

in which the last limit comes from (28) and from the fact that analytic functions, like $\boldsymbol{\nu}\circ\Xi$, have locally bounded derivatives. Moreover, for any $\ell \in \mathcal{I}(\boldsymbol{x}^*)\setminus\mathcal{I}$ or $\ell \in I_q \setminus \mathcal{J}$, it holds

$$
\lim_{k\in\mathcal{K}} \eta_\ell^k = \lim_{k\in\mathcal{K}} \left( \frac{(\boldsymbol{y}^k)^T \nabla A_\ell(\boldsymbol{y}^k)}{A_\ell(\boldsymbol{y}^k)} + \sum_{\mathrm{j}=1}^n \alpha_{\mathrm{j}}^{(\ell)} \right) \overset{(28)}{=} \sum_{\mathrm{j}=1}^n \alpha_{\mathrm{j}}^{(\ell)} \in \mathbb{N}^*. \tag{32}
$$

Equations (31) and (32) allow us to conclude the proof. □

**Remark 4.1.** *From [18, Lemma 5.3.], it is possible to transform any finite set of differentiable subanalytic functions into a set of differentiable analytic functions by composing each function with a proper analytic map. This means that, following the above proof, it is possible to generalize Theorem 4.1 for feasible sets described by differentiable subanalytic functions, which encompass almost all differentiable functions that appear in applications. We have decided not to consider the 'subanalytic context' due to the lack of space to discuss subanalytic sets and functions. With such observation in mind, all the results established here for sets described by analytic functions could be generalized for sets described by differentiable subanalytic functions.*

Aiming at sequential optimality conditions stronger than CAKKT, the limits in (8) were assumed together with further hypotheses in previous works [14, 43]. Here, on the other hand, we will show that the conditions displayed in the implication (**W**) of Definition 3.1 are equivalent to requiring (8), whenever the constraints involved are described by analytic functions. Since the conditions in the implication (**W**) are weaker than (8) in the context of problem (1) (cf. *Discussion* after the proof below), this will help to straighten out the necessary hypotheses to produce a strong sequential optimality condition associated with an algorithm.

Now, we have the necessary tools to demonstrate the main theorem of this article.

***Proof of Theorem 3.1***. Let $\boldsymbol{x}^*$ be a WCAKKT point. Given WCAKKT sequences $\{\boldsymbol{x}^k\}_{k\in\mathbb{N}} \subset \mathbb{R}^n$, and corresponding multipliers $\{\boldsymbol{\mu}^k\}_{k\in\mathbb{N}} \subset \mathbb{R}^p_+$ and $\{\boldsymbol{\lambda}^k\}_{k\in\mathbb{N}} \subset \mathbb{R}^q$, let us suppose, for a proof by *reductio ad absurdum*, that $\{\boldsymbol{x}^k\}_{k\in\mathbb{N}} \subset \mathbb{R}^n$ is not a CAKKT sequence associated with the multipliers $\{\boldsymbol{\mu}^k\}_{k\in\mathbb{N}} \subset \mathbb{R}^p_+$ and $\{\boldsymbol{\lambda}^k\}_{k\in\mathbb{N}} \subset \mathbb{R}^q$, i.e., it does not hold that

$$\lim_{k\to\infty} \mu_i^k c_i(\boldsymbol{x}^k) = 0, \text{ for all } i \in I_p, \quad \lim_{k\to\infty} \lambda_j^k h_j(\boldsymbol{x}^k) = 0, \text{ for all } j \in I_q. \tag{33}$$

Under this hypothesis, for some $\delta > 0$ and $\mathcal{I}_\infty \subset_\infty \mathbb{N}$,

$$\begin{aligned} &\text{there exists } i \in I_p \text{ such that, for all } k \in \mathcal{I}_\infty, |\mu_i^k c_i(\boldsymbol{x}^k)| \geq \delta \\ &\text{or, there exists } j \in I_q \text{ such that, for all } k \in \mathcal{I}_\infty, |\lambda_j^k h_j(\boldsymbol{x}^k)| \geq \delta. \end{aligned} \tag{34}$$

At this point, note that $\{\boldsymbol{x}^k\}_{k\in\mathcal{I}_\infty}$ can be regarded as an AKKT sequence. Hence, Theorem 4.1 ensures the existence of an infinite subset of natural numbers $\mathcal{I}'_\infty \subset_\infty \mathcal{I}_\infty$, and sequences $\{\boldsymbol{\nu}^k\}_{k\in\mathcal{I}'_\infty} \subset \mathbb{R}^p$ and $\{\boldsymbol{\eta}^k\}_{k\in\mathcal{I}'_\infty} \subset \mathbb{R}^q$ so that $\lim_{k\in\mathcal{I}'_\infty} \boldsymbol{\nu}^k = \boldsymbol{\nu} \in (\mathbb{N}^*)^p$, $\lim_{k\in\mathcal{I}'_\infty} \boldsymbol{\eta}^k = \boldsymbol{\eta} \in (\mathbb{N}^*)^q$, and

$$\lim_{k\in\mathcal{I}'_\infty} \left( \textstyle\sum_{i=1}^p \nu_i^k \mu_i^k c_i(\boldsymbol{x}^k) + \sum_{j=1}^q \eta_j^k \lambda_j^k h_j(\boldsymbol{x}^k) \right) = 0. \tag{35}$$

Since the primal sequence $\{\boldsymbol{x}^k\}_{k\in\mathbb{N}}$ is a WCAKKT one, we can take $\mathcal{K}$ equal to $\mathcal{I}'_\infty$ in Definition 3.1 to conclude that, *for all* sequences of real vectors $\{\boldsymbol{\alpha}^k\}_{k\in\mathcal{I}'_\infty} \subset \mathbb{R}^p$ and $\{\boldsymbol{\beta}^k\}_{k\in\mathcal{I}'_\infty} \subset \mathbb{R}^q$ satisfying $\lim_{k\in\mathcal{I}'_\infty} \boldsymbol{\alpha}^k = \boldsymbol{\alpha} \in (\mathbb{N}^*)^p$, $\lim_{k\in\mathcal{I}'_\infty} \boldsymbol{\beta}^k = \boldsymbol{\beta} \in (\mathbb{N}^*)^q$,

$$\begin{aligned} &\text{if} \lim_{k\in\mathcal{I}'_\infty} \left( \textstyle\sum_{i=1}^p \alpha_i^k \mu_i^k c_i(\boldsymbol{x}^k) + \sum_{j=1}^q \beta_j^k \lambda_j^k h_j(\boldsymbol{x}^k) \right) = 0, \text{ necessarily} \\ &\quad \lim_{k\in\mathcal{I}'_\infty} \mu_i^k c_i(\boldsymbol{x}^k) = 0, \text{ for all } i \in I_p, \quad \lim_{k\in\mathcal{I}'_\infty} \lambda_j^k h_j(\boldsymbol{x}^k) = 0, \text{ for all } j \in I_q. \end{aligned}$$

Now, by (35), since the antecedent of the implication above is true for the sequences $\{\boldsymbol{\nu}^k\}_{k\in\mathcal{I}'_\infty} \subset \mathbb{R}^p$ and $\{\boldsymbol{\eta}^k\}_{k\in\mathcal{I}'_\infty} \subset \mathbb{R}^q$, we have

$$\lim_{k\in\mathcal{I}'_\infty} \mu_i^k c_i(\boldsymbol{x}^k) = 0, \text{ for all } i \in I_p, \text{ and } \lim_{k\in\mathcal{I}'_\infty} \lambda_j^k h_j(\boldsymbol{x}^k) = 0, \text{ for all } j \in I_q,$$

which contradicts (34), since $\mathcal{I}'_\infty \subset_\infty \mathcal{I}_\infty$. $\square$

**Discussion** It is worth noticing that in case the functions that describe the constraints of problem (1) are not analytic, the conclusion of Theorem 3.1 may not be valid. In other words, a WCAKKT sequence might not be a CAKKT sequence whenever analyticity does not hold. The counterexample of [13] suffices to reach this conclusion. The particular sequence presented in [13], by construction, is not a CAKKT one. Although the limit point of such a sequence is a CAKKT *point*, due to the existence of other possible sequences converging to the origin which are actually CAKKT sequences, it is worth recalling that the focus of our work is on the sequences and not just on the respective limit points. By not satisfying the Kurdyka-Lojasiewicz inequality, the constraint of the aforementioned counterexample (namely, $h(x) = x^4 \sin(1/x)$, if $x \neq 0$, and $h(0) = 0$) is not analytic [24]. One should observe that in such a case, for an arbitrary sequence of real numbers $\{\eta^k\}_{k\in\mathbb{N}}$ converging to a positive natural number $\eta$ satisfying $\lim_{k\to\infty} \eta^k \lambda^k h(x^k) = 0$, we necessarily have $\lim_{k\to\infty} \lambda^k h(x^k) = \lim_{k\to\infty} (\eta^k)^{-1}(\eta^k \lambda^k h(x^k)) = 0$, since the sequence $\{(\eta^k)^{-1}\}_{k\in\mathbb{N}}$ converges to a positive number $\eta^{-1}$ and $\{\eta^k \lambda^k h(x^k)\}_{k\in\mathbb{N}}$ converges to zero. Hence, such a sequence is indeed a WCAKKT sequence. Further, it is possible to prove that $\{x^k\}_{k\in\mathbb{N}}$ can not be regarded as a CAKKT sequence by any possible choice of multipliers $\{\lambda^k\}_{k\in\mathbb{N}}$. Consequently, our Theorem 3.1 is tight, as its assumptions are essential to reach its conclusions. Moreover, despite implication (**W**) of Definition 3.1 being weaker than (8) in the context of problem (1), these two conditions are equivalent whenever the constraints are stated by analytic functions.

# 5 Algorithmic consequences of the WCAKKT conditions

Concerning optimization methods for which sequences of approximate Lagrange multipliers $\{\boldsymbol{\mu}^k\}_{k\in\mathbb{N}}$ and $\{\boldsymbol{\lambda}^k\}_{k\in\mathbb{N}}$ may be obtained, the corresponding primal sequence $\{\boldsymbol{x}^k\}_{k\in\mathbb{N}}$ may be feasible or infeasible. Among these methods, the ones that fit the WCAKKT conditions are those for which the multiplier estimates and the respective constraints obey the sign agreement below

$$\min\left\{\{\mu_i^k c_i(\boldsymbol{x}^k)\}_{i\in I_p}, \{\lambda_j^k h_j(\boldsymbol{x}^k)\}_{j\in I_q}\right\} \geq 0. \tag{36}$$

Since methods that generate infeasible primal iterates do not necessarily produce Lagrange multiplier approximations with the previous sign agreement, we are driven to the following relaxation

$$\min\left\{\{\mu_i^k c_i(\boldsymbol{x}^k)\}_{i\in I_p}, \{\lambda_j^k h_j(\boldsymbol{x}^k)\}_{j\in I_q}\right\} \geq \theta_k, \tag{37}$$

in which $\{\theta_k\}_{k\in\mathbb{N}} \subset \mathbb{R}$ converges to zero. By verifying such a relaxed sign agreement of the complementarity measure along the iterations, this class of methods has the intrinsic feature of being approximately infeasible concerning the so-called *relevant* inequality constraints of the problem, i.e., those for which the sequence $\{\mu_i^k\} \subset \mathbb{R}_+$ ($i \in I_p$) is bounded away from zero. Therefore, for the sequence generated by a method that obeys (37), the inequality constraints that are relevant to the AKKT stationarity (5) satisfy $c_i(\boldsymbol{x}^k) \geq \theta_k/\mu_i^k$, for $k \in \mathbb{N}$, and large enough. In other words, the relevant constraints are violated by a multiple of $\theta_k$, as shown Figure 2 (left).

On the other hand, not all methods that are compatible with the AKKT conditions generate infeasible primal iterates. For inequality-constrained problems, methods that generate feasible primal iterates usually produce approximate Lagrange multipliers with signs opposite to those of the corresponding constraints. To align such behavior with the WCAKKT conditions, we may consider the following relaxation

$$\max\left\{\{\mu_i^k c_i(\boldsymbol{x}^k)\}_{i\in I_p}, \{\lambda_j^k h_j(\boldsymbol{x}^k)\}_{j\in I_q}\right\} \leq \theta_k, \tag{38}$$

for a scalar sequence $\{\theta_k\}_{k\in\mathbb{N}} \subset \mathbb{R}$ converging to zero. The feasibility of the primal sequence generated by methods for which (38) is satisfied may be noticed for the relevant inequality constraints ($i \in I_p$), since in this case $\mu_i^k > 0$, so that $c_i(\boldsymbol{x}^k) \leq \theta_k/\mu_i^k$, for $k \in \mathbb{N}$, and large enough. Therefore, the relevant constraints are verified up to a multiple of $\theta_k$, as depicted in Figure 2 (right).

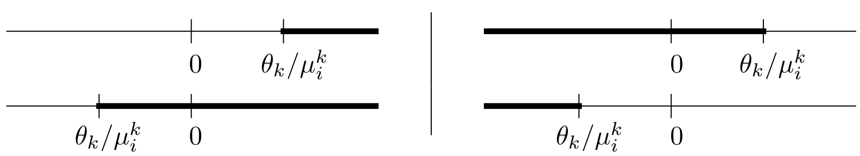


Figure 2: Illustrations for the approximate (in)feasibility of the relevant $i$-th constraint: on the left for an infeasible method satisfying (37), and on the right for a feasible method satisfying (38).

The next result provides a technical tool to be associated with algorithms satisfying either (37) or (38).

**Lemma 5.1.** *Let $s \in \mathbb{N}^*$, $\mathcal{K} \subset_\infty \mathbb{N}$, and $\{\boldsymbol{z}^k\}_{k\in\mathcal{K}} \subset \mathbb{R}^s$ be given, and assume that, for a real-valued scalar sequence $\{\theta_k\}_{k\in\mathcal{K}}$ converging to zero such that for all $i \in I_s$ and $k \in \mathcal{K}$, we have*

(a) *either*

$$z_i^k \geq \theta_k \tag{39}$$

(b) *or*

$$z_i^k \leq \theta_k. \tag{40}$$

*If*

$$\lim_{k\in\mathcal{K}} \textstyle\sum_{i=1}^s \eta_i^k z_i^k = 0, \tag{41}$$

*with $\{\boldsymbol{\eta}^k\}_{k\in\mathcal{K}} \subset \mathbb{R}_+^s$ bounded and away from zero (i.e., there exist $L > 0$ and $\delta > 0$ with $L \geq \eta_i^k \geq \delta > 0$, for all $i \in I_s$ and $k \in \mathcal{K}$), then*

$$\lim_{k\in\mathcal{K}} \boldsymbol{z}^k = \mathbf{0}. \tag{42}$$

*Proof.* First, let us analyze case (a). For simplicity, note that we can assume without losing generality that $\mathcal{K}$ is the entire set $\mathbb{N}$, because all the subsequences may be regarded as a sequence. With this observation in mind, since the limit (42) holds if, and only if, each coordinate of the sequence $\{\boldsymbol{z}^k\}_{k\in\mathbb{N}}$ has limit zero, let us fix some $i \in I_s$ and prove that $\lim_{k\to\infty} z_i^k = 0$. Indeed, saying that $\lim_{k\to\infty} z_i^k = 0$ is equivalent to the statement that, for all $\epsilon > 0$, the set of indexes $k \in \mathbb{N}$ satisfying

$$|z_i^k| > \epsilon. \tag{43}$$

is finite.

By contradiction, let us assume that there exists $\mathcal{I}_\infty \subset_\infty \mathbb{N}$ such that (43) holds for all $k \in \mathcal{I}_\infty$. Making $\theta_k > -\epsilon/2$ for all large enough $k \in \mathcal{I}_\infty$, by (39) and (43), it must occur $z_i^k > \epsilon$, for the same indexes $k \in \mathcal{I}_\infty$. On the other hand, for all large enough $k \in \mathcal{I}_\infty$, we have

$$\begin{aligned}
\textstyle\sum_{j=1}^s \eta_j^k z_j^k &= \eta_i^k z_i^k + \sum_{j=1, j\neq i}^{s} \eta_j^k z_j^k && \left[z_i^k > \epsilon\right] \\
&\geq \delta\epsilon + \sum_{j=1, j\neq i}^{s} \eta_j^k z_j^k && \left[\eta_j^k \geq \delta,\ \forall k \in \mathbb{N},\ \forall j \in I_s\right] \\
&\overset{(39)}{\geq} \delta\epsilon + \min\{\theta_k, 0\}(s-1)L && \left[\eta_j^k \leq L,\ \forall k \in \mathbb{N},\ \forall j \in I_s\right].
\end{aligned}$$

Therefore, taking limits in both sides of the previous inequality for $k \in \mathcal{I}_\infty$, it follows that $\delta\epsilon \leq 0$, in contradiction with the fact that $\delta > 0$ and $\epsilon > 0$. Hence, the set of indexes satisfying (43) is finite, for all $\epsilon > 0$, as desired.

The proof of case (b) comes from applying the previous reasoning to the sequence $\{-\boldsymbol{z}^k\}_{k\in\mathcal{K}}$. □

The following result is fundamental to ensuring the WCAKKT conditions for methods that satisfy either condition (37), like safeguarded augmented Lagrangian methods, inexact restoration methods, infeasible interior-point methods, and the algorithm PACNO within the smooth context, or obey condition (38), such as barrier methods.

**Proposition 5.1.** *Let $\boldsymbol{x}^*$ be a feasible point for problem* (1)*. Suppose that $\boldsymbol{x}^*$ is an AKKT point with the primal sequence $\{\boldsymbol{x}^k\}_{k\in\mathbb{N}} \subset \mathbb{R}^n$ and multipliers $\{\boldsymbol{\mu}^k\}_{k\in\mathbb{N}} \subset \mathbb{R}^p_+$ and $\{\boldsymbol{\lambda}^k\}_{k\in\mathbb{N}} \subset \mathbb{R}^q$. If either* (37) *or* (38) *holds, then $\{\boldsymbol{x}^k\}_{k\in\mathbb{N}}$ is also a WCAKKT sequence with multipliers $\{\boldsymbol{\mu}^k\}_{k\in\mathbb{N}}$ and $\{\boldsymbol{\lambda}^k\}_{k\in\mathbb{N}}$. In algorithmic terms, all methods compatible with AKKT such that either* (37) *or* (38) *holds are also compatible with WCAKKT.*

*Proof.* Let $\boldsymbol{x}^*$ be an AKKT point for problem (1) associated with the primal sequence $\{\boldsymbol{x}^k\}_{k\in\mathbb{N}} \subset \mathbb{R}^n$, and multipliers $\{\boldsymbol{\mu}^k\}_{k\in\mathbb{N}} \subset \mathbb{R}^p_+$ and $\{\boldsymbol{\lambda}^k\}_{k\in\mathbb{N}} \subset \mathbb{R}^q$ , and assume first that (37) holds. Hence, we have the existence of a scalar sequence $\{\theta_k\}_{k\in\mathbb{N}}$ converging to zero such that, for all $k \in \mathbb{N}$,

$$\mu_i^k c_i(\boldsymbol{x}^k) \geq \theta_k, \text{ for all } i \in I_p \text{ and } \lambda_j^k h_j(\boldsymbol{x}^k) \geq \theta_k, \text{ for all } j \in I_q. \tag{44}$$

To show the validity of the WCAKKT conditions, it is sufficient to show that, whenever we start with an arbitrary infinite subset of natural numbers $\mathcal{K}$, and sequences of real vectors $\{\boldsymbol{\nu}^k\}_{k\in\mathcal{K}} \subset \mathbb{R}^p$ and $\{\boldsymbol{\eta}^k\}_{k\in\mathcal{K}} \subset \mathbb{R}^q$ so that $\lim_{k\in\mathcal{K}} \boldsymbol{\nu}^k = \boldsymbol{\nu} \in (\mathbb{N}^*)^p$, $\lim_{k\in\mathcal{K}} \boldsymbol{\eta}^k = \boldsymbol{\eta} \in (\mathbb{N}^*)^q$ and

$$\lim_{k\in\mathcal{K}} \left(\sum_{i=1}^p \nu_i^k \mu_i^k c_i(\boldsymbol{x}^k) + \sum_{j=1}^q \eta_j^k \lambda_j^k h_j(\boldsymbol{x}^k)\right) = 0, \text{ necessarily} \tag{45}$$

$$\lim_{k\in\mathcal{K}} \mu_i^k c_i(\boldsymbol{x}^k) = 0, \text{ for all } i \in I_p, \lim_{k\in\mathcal{K}} \lambda_j^k h_j(\boldsymbol{x}^k) = 0, \text{ for all } j \in I_q. \tag{46}$$

Indeed, for all $k \in \mathcal{K}$, define the sequence $\{\boldsymbol{z}^k\}_{k\in\mathcal{K}}$ with $z_i^k := \mu_i^k c_i(\boldsymbol{x}^k)$, for all $i \in I_p$, and and $z_{j+p}^k := \lambda_j^k h_j(\boldsymbol{x}^k)$, for all $j \in I_q$. Note that inequalities (44) ensure that $z_\ell^k \geq \theta_k$, for all $k \in \mathcal{K}$ and $\ell \in I_{p+q}$. Now, case (a) of Lemma 5.1 guarantee that, whenever (45) holds, i.e.,

$$\lim_{k\in\mathcal{K}} \left(\sum_{i=1}^p \nu_i^k z_i^k + \sum_{j=1}^q \eta_j^k z_{j+p}^k\right) = \lim_{k\in\mathcal{K}} \left(\sum_{i=1}^p \nu_i^k \mu_i^k c_i(\boldsymbol{x}^k) + \sum_{j=1}^q \eta_j^k \lambda_j^k h_j(\boldsymbol{x}^k)\right) = 0,$$

then $\lim_{k\in\mathcal{K}} \boldsymbol{z}^k = \boldsymbol{0}$. In other words, $\lim_{k\in\mathcal{K}} \mu_i^k c_i(\boldsymbol{x}^k) = 0$, for all $i \in I_p$, and $\lim_{k\in\mathcal{K}} \lambda_j^k h_j(\boldsymbol{x}^k) = 0$, for all $j \in I_q$. Thus, $\boldsymbol{x}^*$ is a WCAKKT point.

Now, assuming (38) and applying case (b) of Lemma 5.1, the result is obtained in the same lines as above. □

One of the first results proved for SOCs is that they can be linked with practical methods [1, 8, 13, 33, 39]. The WCAKKT conditions cannot be different. In our scenario, the methods that deserve our initial attention are penalty and barrier methods [17]. The main idea of such algorithms is to convert complex constrained problems into a sequence of simple and unconstrained ones. As a result, constrained problems can be solved by applying techniques of unconstrained optimization. In practice, the sequence of problems does not need to be exactly solved; just a few iterations may be enough for the optimization process to be carried out with success.

To recall the basic process of external penalty algorithms, let us consider a function $P : \mathbb{R}^n \to \mathbb{R}$ that is positive for infeasible points and null for the feasible ones. To transform the constrained problem into a sequence of unconstrained ones, for each iteration $k \in \mathbb{N}$, a positive factor, say $\rho_k$, of the external penalty function $P$ is added to the objective:

$$\min_{\boldsymbol{x}\in\mathbb{R}^n} f(\boldsymbol{x}) + \rho_k P(\boldsymbol{x}). \tag{47}$$

The scalar $\rho_k$ is called penalty parameter. Note that global minimizers of (47) that are feasible for the original constraints are, in fact, global minimizers of the constrained problem. Unfortunately, finding a positive $\rho_k$ for which the global minimizer is feasible for the original constrained problem can only be expected to be accomplished by exact penalty functions, which, in general, are nonsmooth [42]. In the smooth context, algorithmically, to reach both feasibility and optimality, the penalty parameter sequence $\{\rho_k\}_{k\in\mathbb{N}}$ must be driven to infinity, and the sequence of associated problems must be solved exactly [17, Prop. 4.1.1 and 4.2.1]. Since, in general, it is computationally expensive to solve unconstrained problems, in practice inexact minimizations are performed by finding a sequence, say $\{\boldsymbol{x}^k\}_{k\in\mathbb{N}}$, of approximate stationary points such that

$$\nabla f(\boldsymbol{x}^k) + \rho_k \nabla P(\boldsymbol{x}^k) = \boldsymbol{v}^k, \text{ for a sequence } \{\boldsymbol{v}^k\}_{k\in\mathbb{N}} \text{ satisfying } \lim_{k\to\infty} \boldsymbol{v}^k = \mathbf{0}.$$

Under the external penalty methods setting, several $P$ can be considered. For instance, if the function $P$ is taken as $P(\boldsymbol{x}) = 1/2\left(\|\boldsymbol{h}(\boldsymbol{x})\|_2^2 + \|\boldsymbol{c}(\boldsymbol{x})_+\|_2^2\right)$, for all $\boldsymbol{x}\in\mathbb{R}^n$, the associated method is called *quadratic penalty* method; if $P(\boldsymbol{x}) = \|\boldsymbol{h}(\boldsymbol{x})\|_1 + \|\boldsymbol{c}(\boldsymbol{x})_+\|_1$, for all $\boldsymbol{x}\in\mathbb{R}^n$, the associated method is called $\ell_1$ *penalty* method. Note that external penalty functions give a measure for infeasibility, i.e., the larger the value of $P(\boldsymbol{x})$, the farther $\boldsymbol{x}\in\mathbb{R}^n$ is from the feasible set.

Barrier or interior penalty methods are devised for problems with just authentic inequality constraints (feasible set with nonempty interior), and they work with a different strategy. Instead of penalizing the infeasibility, and allowing any point to be reachable, the method adds to the objective function, at each iteration $k\in\mathbb{N}$, a small positive multiple $\gamma_k$ of a continuous map $B:\Omega^\circ\to\mathbb{R}$ that is bounded from below. The map $B$ tends to infinity as a sequence of 'strictly' feasible points approaches the boundary of $\Omega$, being its interior denoted by $\Omega^\circ = \{\boldsymbol{x}\in\mathbb{R} : \boldsymbol{c}(\boldsymbol{x}) < \mathbf{0}\}$. By doing this, the constrained problem is transformed into a new sequence of unconstrained optimization, namely, for each $k\in\mathbb{N}$, it is solved

$$\begin{aligned} &\min_{\boldsymbol{x}\in\mathbb{R}^n} f(\boldsymbol{x}) + \gamma_k B(\boldsymbol{x}) \\ &\text{s.t. } c(\boldsymbol{x}) < \mathbf{0}. \end{aligned} \tag{48}$$

This unconstrained problem creates an obstacle in the boundary of the feasible set, $\Omega$, so methods starting within $\Omega^\circ$ would never approach the boundary at subsequent iterations. For this reason, barrier methods are also known as interior-point methods. Thus, to reach global minimizers, the sequence of positive parameters $\{\gamma_k\}_{k\in\mathbb{N}}$ needs to asymptotically approach zero. By doing this procedure, the cost of reaching the boundary decreases as the iterations advance. As in the external penalty methods, it is required that the unconstrained optimization problem (48) be iteratively globally optimized. Since the expenses of finding the global minimizer are prohibitive, inexact minimizations are usually carried out by finding approximate stationary points:

$$\nabla f(\boldsymbol{x}^k) + \gamma_k \nabla B(\boldsymbol{x}^k) = \boldsymbol{v}^k, \text{ for a sequence } \{\boldsymbol{v}^k\}_{k\in\mathbb{N}} \text{ satisfying } \lim_{k\to\infty} \boldsymbol{v}^k = \mathbf{0},$$

with $c(\boldsymbol{x}^k) < \mathbf{0}$, for each $k\in\mathbb{N}$. For interior-point methods, some choices of the barrier functions $B$ are possible. If the function $B$ is equal to $B(\boldsymbol{x}) = \sum_{i=1}^p -\ln(-c_i(\boldsymbol{x}))$, for all $\boldsymbol{x}\in\mathbb{R}^n$, the associated method is called *log barrier* method, if $B(\boldsymbol{x}) = -\sum_{i=1}^p 1/c_i(\boldsymbol{x})$, for all $\boldsymbol{x}\in\mathbb{R}^n$, the associated method is called *inverse barrier* method [42].

The external penalty method has a rich theory of convergence, despite its numerical drawbacks, being able to generate an AKKT sequence, say $\{\boldsymbol{x}^k\}_{k\in\mathbb{N}}$, and AKKT multipliers [8], say $\{\boldsymbol{\mu}^k\}_{k\in\mathbb{N}}$ and $\{\boldsymbol{\lambda}^k\}_{k\in\mathbb{N}}$. As recently established, the more properties the AKKT multipliers have, the stronger their associated convergence theory is, guaranteeing the KKT conditions under weak constraint qualifications [1, 11, 43]. For the quadratic penalty method, the AKKT multipliers possess the distinguishing feature of having approximate Lagrange multipliers with the same sign as their corresponding constraints

$$\mu_i^k c_i(\boldsymbol{x}^k) \geq 0, \text{ for all } i \in I_p, \text{ and } \lambda_j^k h_j(\boldsymbol{x}^k) \geq 0, \text{ for all } j \in I_q.$$

This gives us an intuition why several external penalty methods fit well into the class of algorithms compatible with the WCAKKT conditions. Considering the relaxations of the above condition, not only external penalty methods fall into this category, but also penalty-barrier methods [25], inexact restoration, safeguarded augmented Lagrangian methods, and the algorithm PACNO in the smooth context.

Similar to the external penalty setting, the log barrier methods for the inequality-constrained case are also related to AKKT sequences [8]. Due to the feasibility of the primal sequence, the signs of the approximate Lagrange multipliers associated with log barrier methods are opposite to the sign of the corresponding constraints, that is $\mu_i^k c_i(\boldsymbol{x}^k) < 0$, for all $i \in I_p$. Due to this important feature, such types of methods are also compatible with the WCAKKT conditions. Moreover, the analysis can be broadened to take into account penalty-barrier approaches (cf. [12]), as allowed by the relaxation stated in expression (38).

In the next result, some essential features of methods compatible with AKKT are shown to ensure compatibility with the WCAKKT conditions as well.

**Theorem 5.1.** *Consider problem* (1)*, and let $\boldsymbol{x}^*$ be one of its feasible points. Assume that the objective function of* (1)*, and all the functions that describe its feasible set are Lipschitz continuous around $\boldsymbol{x}^*$. The WCAKKT conditions hold whenever there exist an AKKT sequence $\{\boldsymbol{x}^k\}_{k\in\mathbb{N}} \subset \mathbb{R}^n$, and multipliers $\{\boldsymbol{\mu}^k\}_{k\in\mathbb{N}} \subset \mathbb{R}^p_+$, $\{\boldsymbol{\lambda}^k\}_{k\in\mathbb{N}} \subset \mathbb{R}^q$ such that $\lim_{k\to\infty} \boldsymbol{x}^k = \boldsymbol{x}^*$, and at least one of the following assumptions hold:*

(A1) $\lim_{k\to\infty} \mu_i^k \min\{c_i(\boldsymbol{x}^k), 0\} = 0$, *for all* $i \in I_p$, *and* $\lim_{k\to\infty} \min\{\lambda_j^k h_j(\boldsymbol{x}^k), 0\} = 0$, *for all* $j \in I_q$*;*

(A2) $\lim_{k\to\infty} \mu_i^k \max\{c_i(\boldsymbol{x}^k), 0\} = 0$, *for all* $i \in I_p$, *and* $\lim_{k\to\infty} \max\{\lambda_j^k h_j(\boldsymbol{x}^k), 0\} = 0$, *for all* $j \in I_q$*;*

(A3) *The function $\boldsymbol{h}$ is identically zero, and there exists a decreasing sequence of positive real numbers $\{\epsilon_k\}_{k\in\mathbb{N}}$ converging to zero so that*

$$\lim_{k\to\infty} \epsilon_k \|\boldsymbol{\mu}^k\|_\infty = 0 \tag{49}$$

*and*

$$\mu_i^k = 0, \text{ for all } i \notin \mathcal{I}_{\epsilon_k}(\boldsymbol{x}^k), \tag{50}$$

*where, for all points* $\boldsymbol{x} \in \mathbb{R}^n$ *and* $\epsilon > 0$,

$$\mathcal{I}_\epsilon(\boldsymbol{x}) := \{i \in I_p : \exists\ \boldsymbol{y} \in \mathcal{B}[\boldsymbol{x}, \epsilon] \text{ with } c_i(\boldsymbol{y}) \geq 0\}.$$

*Proof.* For all cases, let $\{\boldsymbol{x}^k\}_{k\in\mathbb{N}}$ be an AKKT sequence associated with the multipliers $\{\boldsymbol{\mu}^k\}_{k\in\mathbb{N}} \subset \mathbb{R}^p_+$, $\{\boldsymbol{\lambda}^k\}_{k\in\mathbb{N}} \subset \mathbb{R}^q$ of the respective statement.

Under the assumption of (A1), since $\mu_i^k \geq 0$, for all $i \in I_p$, note that

$$\mu_i^k c_i(\boldsymbol{x}^k) \geq \mu_i^k \min\{c_i(\boldsymbol{x}^k), 0\}, \text{ for all } i \in I_p, \tag{51}$$

$$\text{and } \lambda_j^k h_j(\boldsymbol{x}^k) \geq \min\{\lambda_j^k h_j(\boldsymbol{x}^k), 0\}, \text{ for all } j \in I_q. \tag{52}$$

Now, from the hypotheses, $\lim_{k\to\infty} \mu_i^k \min\{c_i(\boldsymbol{x}^k), 0\} = 0$, for all $j \in I_p$, and $\lim_{k\to\infty} \min\{\lambda_j^k h_j(\boldsymbol{x}^k), 0\} = 0$, for all $i \in I_q$. Hence, taking

$$\epsilon_k := -\left(\sum_{l=1}^p \mu_l^k \min\{c_l(\boldsymbol{x}^k), 0\} + \sum_{j=1}^q \min\{\lambda_j^k h_j(\boldsymbol{x}^k), 0\}\right),$$

by (51) and (52), we have

$$\mu_i^k c_i(\boldsymbol{x}^k) \geq -\epsilon_k, \text{ for all } i \in I_p, \tag{53}$$

$$\text{and } \lambda_j^k h_j(\boldsymbol{x}^k) \geq -\epsilon_k, \text{ for all } j \in I_q, \tag{54}$$

with $\lim_{k\to\infty} \epsilon_k = 0$. Proposition 5.1 let us conclude that $\boldsymbol{x}^*$ is a WCAKKT point with the WCAKKT sequence $\{\boldsymbol{x}^k\}_{k\in\mathbb{N}}$, and WCAKKT multipliers $\{\boldsymbol{\mu}^k\}_{k\in\mathbb{N}} \subset \mathbb{R}^p_+$, $\{\boldsymbol{\lambda}^k\}_{k\in\mathbb{N}} \subset \mathbb{R}^q$.

Assuming (A2), the proof proceeds as previously. Indeed, note that

$$\mu_i^k c_i(\boldsymbol{x}^k) \leq \mu_i^k \max\{c_i(\boldsymbol{x}^k), 0\}, \text{ for all } i \in I_p$$

$$\text{and } \lambda_j^k h_j(\boldsymbol{x}^k) \leq \max\{\lambda_j^k h_j(\boldsymbol{x}^k), 0\}, \text{ for all } j \in I_q.$$

Moreover, setting $\epsilon_k := \sum_{l=1}^p \mu_l^k \max\{c_l(\boldsymbol{x}^k), 0\} + \sum_{i=1}^q \max\{\lambda_j^k h_j(\boldsymbol{x}^k), 0\}$, we have $\mu_i^k c_i(\boldsymbol{x}^k) \leq \epsilon_k$, for all $i \in I_p$, and $\lambda_j^k h_j(\boldsymbol{x}^k) \leq \epsilon_k$, for all $j \in I_q$, with $\lim_{k\to\infty} \epsilon_k = 0$. From Proposition 5.1, it follows that $\boldsymbol{x}^*$ is a WCAKKT point with the WCAKKT sequence $\{\boldsymbol{x}^k\}_{k\in\mathbb{N}}$ and WCAKKT multipliers $\{\boldsymbol{\mu}^k\}_{k\in\mathbb{N}} \subset \mathbb{R}^p_+$, $\{\boldsymbol{\lambda}^k\}_{k\in\mathbb{N}} \subset \mathbb{R}^q$.

Now, under the assumption of (A3), let us define $\vartheta_k := \max_{i\in I_p} L_{c_i} \epsilon_k \|\boldsymbol{\mu}^k\|_\infty$, being $L_{c_i}$ the Lipschitz constant of the $i$-th component of $\boldsymbol{c}$ around $\boldsymbol{x}^*$. Note that, due to (49), $\lim_{k\to\infty} \vartheta_k = 0$.

Two cases should be considered: $i \in \mathcal{I}_{\epsilon_k}(\boldsymbol{x}^k)$ and $i \notin \mathcal{I}_{\epsilon_k}(\boldsymbol{x}^k)$. For the first one, the definition of $\mathcal{I}_{\epsilon_k}(\boldsymbol{x}^k)$ ensures that there exists some $\boldsymbol{x}_i^k \in \mathcal{B}[\boldsymbol{x}^k, \epsilon_k]$ so that $c_i(\boldsymbol{x}_i^k) \geq 0$. This means that

$$\begin{aligned}
\mu_i^k c_i(\boldsymbol{x}^k) &= \mu_i^k (c_i(\boldsymbol{x}^k) - c_i(\boldsymbol{x}_i^k)) + \mu_i^k c_i(\boldsymbol{x}_i^k) \\
&\geq \mu_i^k \left(c_i(\boldsymbol{x}^k) - c_i(\boldsymbol{x}_i^k)\right) \\
&\geq\ - L_{c_i} \mu_i^k \|\boldsymbol{x}^k - \boldsymbol{x}_i^k\| \\
&\geq\ - \max_{i\in I_p} L_{c_i} \mu_i^k \epsilon_k
\end{aligned}$$

$$\geq -\max_{i \in I_p} L_{c_i} \|\boldsymbol{\mu}^k\|_\infty \epsilon_k.$$
$$= -\vartheta_k,$$

for all $k$ large enough.

On the other hand, if $i \notin \mathcal{I}(\boldsymbol{x}^k)$, we have that $\mu_i^k = 0$. Hence, $\mu_i^k c_i(\boldsymbol{x}^k) = 0 \geq -\vartheta_k$. In both cases, for all $k \in \mathbb{N}$ and $i \in I_p$, we have

$$\mu_i^k c_i(\boldsymbol{x}^k) \geq -\vartheta_k. \tag{55}$$

Since $\boldsymbol{h}$ and its Jacobian are both identically zero, Proposition 5.1 ensures that $\boldsymbol{x}^*$ is a WCAKKT point with the WCAKKT sequence $\{\boldsymbol{x}^k\}_{k\in\mathbb{N}}$ and WCAKKT multipliers $\{\boldsymbol{\mu}^k\}_{k\in\mathbb{N}} \subset \mathbb{R}^p_+$, $\{\boldsymbol{\lambda}^k\}_{k\in\mathbb{N}} \subset \mathbb{R}^q$. □

**Remark 5.1.** *Under the assumptions of Theorem 5.1, in particular, the applicability of the WCAKKT conditions may be appreciated as follows:*

- *For problems described only by inequality constraints, and under the same hypotheses of [29, Thm. 3.9], there exists an infinite set $\mathcal{K} \subset_\infty \mathbb{N}$, such that every convergent subsequence $\{\boldsymbol{x}^k\}_{k\in\mathcal{K}}$ generated by the inexact restoration method can be regarded as a WCAKKT sequence, since, considering [14, Thm. 2.7], such a sequence fulfills Assumption* (A1)*.*
- *Assumption* (A2) *says that any feasible method compatible with the AKKT conditions ensures the WCAKKT conditions. In particular, considering problems described only by inequalities, the feasible AKKT sequence $\{\boldsymbol{x}^k\}_{\in\mathbb{N}}$ with limit point $\boldsymbol{x}^*$ generated by computing approximate stationary points of the inverse barrier method [17, 30] is a WCAKKT sequence.*
- *If the function $\boldsymbol{h}$ is identically zero, i.e., if the problem is described only by smooth inequality constraints, every limit point $\boldsymbol{x}^*$ of the algorithm PACNO [32, 33], under the hypotheses of [33, Thm. 4.1(b)], is a WCAKKT point since, in such a context, PACNO generates a sequence satisfying Assumption* (A3)*.*

The next result provides the grounds upon which penalty-barrier [25] and safeguarded augmented Lagrangian [19] methods may be seen as compatible with the WCAKKT conditions.

**Theorem 5.2.** *Consider problem* (1)*, and let $\boldsymbol{x}^*$ be one if its feasible points. $\boldsymbol{x}^*$ is a WCAKKT point whenever there exist a sequence $\{\boldsymbol{x}^k\}_{k\in\mathbb{N}} \subset \mathbb{R}^n$ converging to $\boldsymbol{x}^*$, unbounded and increasing sequences $\{\rho_k\}_{k\in\mathbb{N}} \subset \mathbb{R}^*_+$ and $\{\sigma_k\}_{k\in\mathbb{N}} \subset \mathbb{R}^*_+$ and a sequence $\{\boldsymbol{v}^k\}_{k\in\mathbb{N}} \subset \mathbb{R}^n$ such that $\lim_{k\to\infty} \boldsymbol{v}^k = \boldsymbol{0}$ and at least one of the following hypotheses hold*

(H1) *For all $k \in \mathbb{N}$, we have $c_i(\boldsymbol{x}^k) < 0$ and*

$$\nabla f(\boldsymbol{x}^k) - \textstyle\sum_{i=1}^p \frac{1}{\rho_k c_i(\boldsymbol{x}^k)} \nabla c_i(\boldsymbol{x}^k) + \sum_{j=1}^q \sigma_k h_j(\boldsymbol{x}^k) \nabla h_j(\boldsymbol{x}^k) = \boldsymbol{v}^k.$$

(H2) *For bounded sequences* $\{\bar{\boldsymbol{\mu}}^k\}_{k\in\mathbb{N}} \subset \mathbb{R}^p_+$ *and* $\{\bar{\boldsymbol{\lambda}}^k\}_{k\in\mathbb{N}} \subset \mathbb{R}^q$, *for all* $k \in \mathbb{N}$,

$$\nabla f(\boldsymbol{x}^k) + \textstyle\sum_{i=1}^p \left(\bar{\mu}_i^k + \rho_k c_i(\boldsymbol{x}^k)\right)_+ \nabla c_i(\boldsymbol{x}^k) + \\ \textstyle\sum_{j=1}^q \left(\bar{\lambda}_j^k + \rho_k h_i(\boldsymbol{x}^k)\right) \nabla h_i(\boldsymbol{x}^k) = \boldsymbol{v}^k.$$

*Proof.* Assume first that (H1) holds. Defining $\mu_i^k := -1/(\rho_k c_i(\boldsymbol{x}^k))$, for all $i \in I_p$, and $\lambda_j := \sigma_k h_j(\boldsymbol{x}^k)$, for all $j \in I_q$, from the hypotheses we have

$$\nabla f(\boldsymbol{x}^k) + \textstyle\sum_{i=1}^p \mu_i^k \nabla c_i(\boldsymbol{x}^k) + \sum_{j=1}^q \lambda_j^k \nabla h_j(\boldsymbol{x}^k) = \boldsymbol{v}^k, \tag{56}$$

with $\mu_i^k \geq 0$, $\mu_i^k c_i(\boldsymbol{x}^j) < 0$, for all $i \in I_p$, $\lambda_j^k h_j(\boldsymbol{x}^k) \geq 0$, for all $j \in I_q$, and every $k \in \mathbb{N}$. Hence, for all $i \in I_p$ and $j \in I_q$,

$$\mu_i^k c_i(\boldsymbol{x}^k) \geq \min_{i\in I_p} \mu_i^k c_i(\boldsymbol{x}^k) \ \text{ and } \ \lambda_j^k h_j(\boldsymbol{x}^k) \geq \min_{i\in I_p} \mu_i^k c_i(\boldsymbol{x}^k). \tag{57}$$

Observe that $\lim_{k\to\infty} \mu_i^k c_i(\boldsymbol{x}^k) = \lim_{k\to\infty} -1/\rho_k = 0$, for all $i \in I_p$. This means that the right-hand side of both inequalities in (57) converges to zero. Now, note that, if $c_i(\boldsymbol{x}^*) < 0$, necessarily the non-negative sequence of real numbers $\{\mu_i^k\}_{k\in\mathbb{N}}$ converges to zero, otherwise $\mu_i^k c_i(\boldsymbol{x}^k) < -\delta$, for an infinite set of indexes $k$ and a real number $\delta > 0$. This is a contradiction with (57). Hence, AKKT conditions hold.

Further, defining $\{\epsilon_k\}_{k\in\mathbb{N}}$ by

$$\epsilon_k := \min_{i\in I_p} \mu_i^k c_i(\boldsymbol{x}^k), \tag{58}$$

we have $\lim_{k\to\infty} \epsilon_k = 0$, and by (57), we obtain

$$\mu_i^k c_i(\boldsymbol{x}^k) \geq \theta_k \text{ and } \lambda_j^k h_j(\boldsymbol{x}^k) \geq \theta_k. \tag{59}$$

Due to (56) and (59), Proposition 5.1 ensures that $\boldsymbol{x}^*$ is a WCAKKT point with the WCAKKT sequence $\{\boldsymbol{x}^k\}_{k\in\mathbb{N}}$ and WCAKKT multipliers $\{\boldsymbol{\mu}^k\}_{k\in\mathbb{N}} \subset \mathbb{R}^p_+$, $\{\boldsymbol{\lambda}^k\}_{k\in\mathbb{N}} \subset \mathbb{R}^q$.

Now, assume that (H2) holds. Defining, for all $k \in \mathbb{N}$,

$$\mu_i^k := \left(\bar{\mu}_i^k + \rho_k c_i(\boldsymbol{x}^k)\right)_+, \text{ for all } i \in I_p, \tag{60}$$

$$\text{and } \lambda_j^k := \left(\bar{\lambda}_j^k + \rho_k h_j(\boldsymbol{x}^k)\right), \text{ for all } j \in I_q, \tag{61}$$

from the hypotheses, we can guarantee that

$$\nabla f(\boldsymbol{x}^k) + \textstyle\sum_{i=1}^p \mu_i^k \nabla c_i(\boldsymbol{x}^k) + \sum_{j=1}^q \lambda_j^k \nabla h_j(\boldsymbol{x}^k) = \boldsymbol{v}^k, \tag{62}$$

for all $k \in \mathbb{N}$. Note that, for all $i \in I_p$,

$$\begin{aligned} \mu_i^k c_i(\boldsymbol{x}^k) &= \left(\left(\bar{\mu}_i^k + \rho_k c_i(\boldsymbol{x}^k)\right)_+ - \rho_k c_i(\boldsymbol{x}^k)_+\right) c_i(\boldsymbol{x}^k) + \rho_k c_i(\boldsymbol{x}^k)_+ c_i(\boldsymbol{x}^k) \\ &\geq \left(\left(\bar{\mu}_i^k + \rho_k c_i(\boldsymbol{x}^k)\right)_+ - \rho_k c_i(\boldsymbol{x}^k)_+\right) c_i(\boldsymbol{x}^k) \\ &= \tilde{\mu}_i^k c_i(\boldsymbol{x}^k), \end{aligned}$$

with $\tilde{\mu}_i^k := \left(\bar{\mu}_i^k + \rho_k c_i(\boldsymbol{x}^k)\right)_+ - \rho_k c_i(\boldsymbol{x}^k)_+$. Now, for each $i \in I_p$, we have $\lim_{k\to\infty} \tilde{\mu}_i^k c_i(\boldsymbol{x}^k) = 0$. Indeed, $\{\tilde{\mu}_i^k\}_{k\in\mathbb{N}}$ is bounded, since the scalar function $(\cdot)_+$ is non-expansive and positively homogeneous. Moreover, whenever $c_i(\boldsymbol{x}^*) < 0$, we have $\tilde{\mu}_i^k = 0$, for $k$ large enough, because $\lim_{k\to\infty} \rho_k c_i(\boldsymbol{x}^k) + \bar{\mu}_i^k = -\infty$. Therefore, $\lim_{k\to\infty} \tilde{\mu}_i^k c_i(\boldsymbol{x}^k) = 0$ holds true either if $c_i(\boldsymbol{x}^*) < 0$ or $c_i(\boldsymbol{x}^*) = 0$.

On the other hand, for each $j \in I_q$,

$$\lambda_j^k h_j(\boldsymbol{x}^k) = \left(\bar{\lambda}_j^k + \rho_k h_j(\boldsymbol{x}^k)\right) h_j(\boldsymbol{x}^k) \geq \bar{\lambda}_j^k h_j(\boldsymbol{x}^k),$$

with $\lim_{k\to\infty} \bar{\lambda}_j^k h_j(\boldsymbol{x}^k) = 0$, since $\{\bar{\lambda}_j^k\}_{j\in\mathbb{N}}$ is bounded.

The previous reasoning ensures the existence of a sequence of positive real numbers, say $\{\epsilon_k\}_{k\in\mathbb{N}}$, converging to zero, and satisfying, for all $k \in \mathbb{N}$,

$$\mu_i^k c_i(\boldsymbol{x}^k) \geq \epsilon_k, \text{ for all } i \in I_p \text{ and } \lambda_j^k h_j(\boldsymbol{x}^k) \geq \epsilon_k, \text{ for all } j \in I_q. \tag{63}$$

The inequalities (63) help us to conclude that, if $c_i(\boldsymbol{x}^*) < 0$, the non-negative sequence $\{\mu_i^k\}_{k\in\mathbb{N}}$ cannot have a subsequence bounded away from zero, otherwise $c_i(\boldsymbol{x}^k)\mu_i^k < -\delta$, for a positive scalar $\delta$ and infinitely many values of the index $k$. Hence, $\lim_{k\to\infty} \mu_i^k = 0$. Thus, the sequence $\{\boldsymbol{x}^k\}_{k\in\mathbb{N}}$ is an AKKT sequence.

Now, due to the expressions (62) and (63), Proposition 5.1 ensures that $\boldsymbol{x}^*$ is a WCAKKT point with the WCAKKT sequence $\{\boldsymbol{x}^k\}_{k\in\mathbb{N}}$ and WCAKKT multipliers $\{\boldsymbol{\mu}^k\}_{k\in\mathbb{N}} \subset \mathbb{R}^p_+$, $\{\boldsymbol{\lambda}^k\}_{k\in\mathbb{N}} \subset \mathbb{R}^q$. □

**Remark 5.2.** *For AKKT sequences and corresponding multipliers, under the hypotheses of Theorem 5.2, the applicability of the WCAKKT conditions may be recognized as follows:*

- *Due to* (H1)*, the infeasible interior-point method of [12, Algorithm 1] generates sequences such that every convergent subsequence may be regarded as a WCAKKT sequence. The fact that the method of [12] indeed generates CAKKT sequences is a consequence of the analysis presented in Remark 5.3.*

- *Regarding interior-point methods distinct from the Algorithm 1 of [12], a few words are in order. In the recent work [27], the authors deal with the inequality-constrained nonlinear programming problem, in which slack variables are inserted. They also present a sequential optimality condition that, despite being also nominated as CAKKT, it is distinct and weaker than the original CAKKT condition as proposed by [13]. A detailed discussion along this line is provided on pages 15–16 and Example 3.3 of [14].*

- *Under* (H2)*, the safeguarded augmented Lagrangian method [13, Algorithm 5.1.] generates sequences such that every subsequence converging to $\boldsymbol{x}^*$ may be regarded as a WCAKKT sequence, whenever the sequence of penalty parameters is unbounded. On the other hand, it is not hard to prove that AKKT sequences associated with bounded multipliers are also WCAKKT sequences, assuming* (A1) *or* (A2) *of Theorem 5.1. Now,* *since the augmented Lagrangian method generates AKKT sequences and AKKT multipliers which are bounded whenever the penalty parameter is bounded, the safeguarded augmented Lagrangian method always generates WCAKKT sequences.*

**Remark 5.3.** *As considered in [12], it is usual to apply interior-point algorithms to problems of the form* $\min_{\hat{\boldsymbol{x}}\in\mathbb{R}^m} \bar{f}(\hat{\boldsymbol{x}})$ *s.t.* $\bar{\boldsymbol{h}}(\hat{\boldsymbol{x}}) = \mathbf{0},\ \hat{\boldsymbol{x}} \geq \mathbf{0}$. *To see problem* (1) *in such a form it is enough to introduce slacks and express its original free-of-sign variables as differences of non-negative variables, thus obtaining*

$$\begin{array}{cc} \min\limits_{(\boldsymbol{x},\boldsymbol{y},\boldsymbol{z})\in\mathbb{R}^{2n+p}} & f(\boldsymbol{x}-\boldsymbol{y}) \\ s.t. & \boldsymbol{c}(\boldsymbol{x}-\boldsymbol{y})+\boldsymbol{z}=\mathbf{0} \\ & \boldsymbol{h}(\boldsymbol{x}-\boldsymbol{y})=\mathbf{0} \\ & -\boldsymbol{x}\leq\mathbf{0},\ -\boldsymbol{y}\leq\mathbf{0},\ -\boldsymbol{z}\leq\mathbf{0}. \end{array} \tag{64}$$

*The relationships between CAKKT points of problem* (64) *and CAKKT points of problem* (1) *are established next. Suppose* $(\boldsymbol{x}^*, \boldsymbol{y}^*, \boldsymbol{z}^*)$ *is a CAKKT point of* (64). *Therefore, there exist CAKKT sequences* $\{\boldsymbol{x}^k\}_{k\in\mathbb{N}}$, $\{\boldsymbol{y}^k\}_{k\in\mathbb{N}}$ *and* $\{\boldsymbol{z}^k\}_{k\in\mathbb{N}}$ *converging to* $\boldsymbol{x}^*$, $\boldsymbol{y}^*$ *and* $\boldsymbol{z}^*$, *respectively, and CAKKT multipliers* $\{\boldsymbol{\alpha}^k\}_{k\in\mathbb{N}} \subset \mathbb{R}^n_+$, $\{\boldsymbol{\beta}^k\}_{k\in\mathbb{N}} \subset \mathbb{R}^n_+$, $\{\boldsymbol{\omega}^k\}_{k\in\mathbb{N}} \subset \mathbb{R}^p$, $\{\boldsymbol{\mu}^k\}_{k\in\mathbb{N}} \subset \mathbb{R}^p_+$ *and* $\{\boldsymbol{\lambda}^k\}_{k\in\mathbb{N}} \subset \mathbb{R}^q$ *satisfying the following conditions:*

$$\nabla f(\boldsymbol{x}^k-\boldsymbol{y}^k)+\sum_{i=1}^{p}\omega_i^k\nabla c_i(\boldsymbol{x}^k-\boldsymbol{y}^k)+\sum_{j=1}^{q}\lambda_j^k\nabla h_j(\boldsymbol{x}^k-\boldsymbol{y}^k)-\boldsymbol{\alpha}^k=\boldsymbol{p}^k \tag{65}$$

$$\nabla f(\boldsymbol{x}^k-\boldsymbol{y}^k)+\sum_{i=1}^{p}\omega_i^k\nabla c_i(\boldsymbol{x}^k-\boldsymbol{y}^k)+\sum_{j=1}^{q}\lambda_j^k\nabla h_j(\boldsymbol{x}^k-\boldsymbol{y}^k)+\boldsymbol{\beta}^k=\boldsymbol{q}^k \tag{66}$$

$$\boldsymbol{\omega}^k-\boldsymbol{\mu}^k=\boldsymbol{r}^k \tag{67}$$

$$\omega_i(c_i(\boldsymbol{x}^k-\boldsymbol{y}^k)+z_i^k)=s_i^k,\ \textit{for all } i\in I_p \tag{68}$$

$$\lambda_j^k h_j(\boldsymbol{x}^k-\boldsymbol{y}^k)=t_j^k,\ \textit{for all } i\in I_q \tag{69}$$

$$\alpha_\ell^k x_\ell^k=u_\ell^k,\quad \beta_\ell^k y_\ell^k=v_\ell^k,\quad \textit{and}\quad \mu_i^k z_i^k=w_i^k,\ \textit{for all } \ell\in I_n \textit{ and } i\in I_p, \tag{70}$$

*for all* $k \in \mathbb{N}$, *in which the sequences on the right-hand sides have appropriate dimensions, and all of them converge to zero. Observe that equations* (65)–(67) *are related to the limit* (5), *whereas equations* (68)–(70) *are related to the limits* (8).

*Subtracting equation* (65) *from* (66), *we obtain* $\boldsymbol{q}^k - \boldsymbol{p}^k = \boldsymbol{\beta}^k + \boldsymbol{\alpha}^k$ *for all* $k \in \mathbb{N}$. *Since the sequences* $\{\boldsymbol{p}^k\}_{k\in\mathbb{N}}$ *and* $\{\boldsymbol{q}^k\}_{k\in\mathbb{N}}$ *converge to zero, applying Lemma 5.1, the sequences of positive terms* $\{\boldsymbol{\beta}^k\}$ *and* $\{\boldsymbol{\alpha}^k\}_{k\in\mathbb{N}}$ *also converge to* *zero. Now, the substitution of the relationship* (67) *into* (65) *gives that*

$$\nabla f(\boldsymbol{x}^k - \boldsymbol{y}^k) + \sum_{i=1}^{p} \mu_i^k \nabla c_i(\boldsymbol{x}^k - \boldsymbol{y}^k) + \sum_{i=1}^{p} \lambda_j^k \nabla h_j(\boldsymbol{x}^k - \boldsymbol{y}^k) = $$
$$= \boldsymbol{p}^k - \sum_{i=1}^{p} r_i^k \nabla c_i(\boldsymbol{x}^k - \boldsymbol{y}^k) + \boldsymbol{\alpha}^k,$$

*for all $k \in \mathbb{N}$. Hence, since the limit of the right-hand side above is zero, the limit* (5) *holds.*

*Additionally, making the substitution of equation* (67) *into* (68), *we have*

$$\begin{aligned} \mu_i^k c_i(\boldsymbol{x}^k - \boldsymbol{y}^k) &= s_i^k - \mu_i^k z_i^k - r_i^k(c_i(\boldsymbol{x}^k - \boldsymbol{y}^k) + z_i^k) \\ &\overset{(70)}{=} v_i^k - w_i^k - r_i^k(c_i(\boldsymbol{x}^k - \boldsymbol{y}^k) + z_i^k). \end{aligned}$$

*Again, the right-hand side of the previous equation goes to zero. Thus, such a limit, together with* (69), *implies the validity of the limits in* (8).

*Finally, note that the derived equations allow us to conclude that the CAKKT conditions are valid at the feasible point $\boldsymbol{x}^* - \boldsymbol{y}^*$ of* (1), *when taking into account the CAKKT sequence $\{\boldsymbol{x}^k - \boldsymbol{y}^k\}_{k\in\mathbb{N}} \subset \mathbb{R}^n$ and the CAKKT multipliers $\{\boldsymbol{\lambda}^k\}_{k\in\mathbb{N}} \subset \mathbb{R}^q$ and $\{\boldsymbol{\mu}^k\}_{k\in\mathbb{N}} \subset \mathbb{R}^p_+$.*

*Conversely, assume that the CAKKT conditions for problem* (1) *hold with the CAKKT sequence $\{\bar{\boldsymbol{x}}^k\}_{k\in\mathbb{N}} \subset \mathbb{R}^n$ and the CAKKT multipliers $\{\boldsymbol{\lambda}^k\}_{k\in\mathbb{N}} \subset \mathbb{R}^q$ and $\{\boldsymbol{\mu}^k\}_{k\in\mathbb{N}} \subset \mathbb{R}^p_+$. Setting the sequences $\{\boldsymbol{x}^k\}_{k\in\mathbb{N}}$, $\{\boldsymbol{y}^k\}$ and $\{\boldsymbol{z}^k\}_{k\in\mathbb{N}}$, respectively, as $\{\bar{\boldsymbol{x}}^k_+\}_{k\in\mathbb{N}}$, $\{-\bar{\boldsymbol{x}}^k_-\}_{k\in\mathbb{N}}$ and $\{-\boldsymbol{c}(\bar{\boldsymbol{x}}^k)\}_{k\in\mathbb{N}}$, define the sequences $\{\boldsymbol{\alpha}^k\}_{k\in\mathbb{N}}$ and $\{\boldsymbol{\beta}^k\}_{k\in\mathbb{N}}$ with identically null terms, and set $\{\boldsymbol{\omega}^k\}_{k\in\mathbb{N}} = \{\boldsymbol{\mu}^k\}_{k\in\mathbb{N}}$. The remaining sequences that must be defined to obtain* (65)–(70) *can be identified trivially. Hence, the validity of the CAKKT conditions for problem* (64) *follows.*

So far, we have pointed out several methods that generate WCAKKT sequences. Hence, we dare to affirm that the WCAKKT conditions are general enough to fit within the convergence analysis of generic barrier and penalty algorithms. Also, since globalized SQP methods employ penalty functions as merit functions, we further conjecture that such methods, like [28], can generate WCAKKT sequences. Due to its wide applicability, the new condition may help to provide strong global convergence results for a wide range of algorithms. Finally, one should notice that, due to Theorem 3.1 and Proposition 3.1, under the analyticity of the model functions that algebraically describe problem (1), a method can only surely attain CAKKT sequences by guaranteeing the generation of WCAKKT sequences.

At this point, all the expected results have been derived, but we are still left with the question raised in Section 3: what are the connections that one needs to identify to ensure that CAKKT is implied by AGP? For a reader familiar with inexact restoration methods, the question has already been answered. Formulating the AGP conditions to obtain multipliers, the conditions can be regarded as satisfying assumption (A2) in the inequality-constrained case. Thus, when the problem is described by inequality constraints and by analytic functions, the primal AGP sequence is also a CAKKT one, with the WCAKKT conditions serving as a link between these two conditions. Additionally, the *strong AGP* (SAGP) condition can be understood as the AGP conditions for optimization problems with just inequality constraints (any equality within the feasible set description is replaced by two inequalities) [40]. Hence, the SAGP sequences can be regarded as CAKKT sequences. The new result derived is surprising, since, for the equality-constrained case, the pure AGP conditions are equivalent to the AKKT conditions – see [14, Thm. 2.7] –, which is the weakest sequential optimality condition known. Now, since this matter motivated our study, we provide a complete statement and detailed proof.

**Proposition 5.2.** *Under the same hypotheses of Theorem 3.1, the SAGP primal sequences are exactly the CAKKT ones. In particular, SAGP points coincide with CAKKT points.*

*Proof.* Let us recall an equivalence result before starting the proof. As previously mentioned, the SAGP conditions are equivalent to the AGP optimality conditions applied to the feasible set given by the inequalities: $\boldsymbol{g}(\boldsymbol{x}) \leq \mathbf{0}$, $\boldsymbol{h}(\boldsymbol{x}) \leq \mathbf{0}$ and $-\boldsymbol{h}(\boldsymbol{x}) \leq \mathbf{0}$, see discussion before the SAGP definition in [40]. By [14, Thm. 2.7], this means that the SAGP conditions are equivalent to the existence of multipliers $\{\boldsymbol{\mu}^k\}_{k\in\mathbb{N}} \subset \mathbb{R}^p_+$, $\{\boldsymbol{\lambda}^k_{(-)}\}_{k\in\mathbb{N}} \subset \mathbb{R}^q_+$ and $\{\boldsymbol{\lambda}^k_{(+)}\}_{k\in\mathbb{N}} \subset \mathbb{R}^q_+$, and a sequence $\{\boldsymbol{x}^k\}_{k\in\mathbb{N}}$ such that $\boldsymbol{x}^k \to \boldsymbol{x}^*$,

$$\lim_{k\to\infty} \Big[\nabla f(\boldsymbol{x}^k) + \textstyle\sum_{i=1}^p \mu_i^k \nabla c_i(\boldsymbol{x}^k) \; + \\ \textstyle\sum_{j=1}^q \Big((\lambda^k_{(+)})_j - (\lambda^k_{(-)})_j\Big) \nabla h_j(\boldsymbol{x}^k)\Big] = \mathbf{0}, \quad (71)$$

$$\lim_{k\to\infty} (\lambda^k_{(+)})_j \min\{h_j(\boldsymbol{x}^k), 0\} = 0, \quad (72)$$

$$\lim_{k\to\infty} (\lambda^k_{(-)})_j \min\{-h_j(\boldsymbol{x}^k), 0\} = 0 \text{ and} \quad (73)$$

$$\lim_{k\to\infty} \mu_i^k \min\{c_i(\boldsymbol{x}^k), 0\} = 0. \quad (74)$$

For the first part, it is enough to prove that the SAGP sequence verifies the WCAKKT conditions, since the WCAKKT sequences are exactly the CAKKT ones, by Theorem 3.1. Under the assumption that $\boldsymbol{x}^*$ satisfies the SAGP conditions, we have the existence of multipliers $\{\boldsymbol{\mu}^k\}_{k\in\mathbb{N}} \subset \mathbb{R}^p_+$, $\{\boldsymbol{\lambda}^k_{(-)}\}_{k\in\mathbb{N}} \subset \mathbb{R}^q_+$ and $\{\boldsymbol{\lambda}^k_{(+)}\}_{k\in\mathbb{N}} \subset \mathbb{R}^q_+$ and a sequence $\{\boldsymbol{x}^k\}_{k\in\mathbb{N}}$ satisfying the limits (71)–(74), with $\{\boldsymbol{x}^k\}_{k\in\mathbb{N}}$ converging to $\boldsymbol{x}^*$.

Moreover, for all $k \in \mathbb{N}$, $i \in I_p$ and $j \in I_q$, it holds

$$\Big((\lambda^k_{(+)})_j - (\lambda^k_{(-)})_j\Big)\, h_j(\boldsymbol{x}^k) = (\lambda^k_{(+)})_j h_j(\boldsymbol{x}^k) - (\lambda^k_{(-)})_j h_j(\boldsymbol{x}^k) \\ \geq (\lambda^k_{(+)})_j \min\{h_j(\boldsymbol{x}^k), 0\} + (\lambda^k_{(-)})_j \min\{-h_j(\boldsymbol{x}^k), 0\}$$

and $\mu_i^k c_i(\boldsymbol{x}^k) \geq \mu_i^k \min\{c_i(\boldsymbol{x}^k), 0\}$. Note that the right-hand sides of the last two inequalities converge to zero, by the limits (72)–(74). Hence, we can take a sequence converging to zero, say $\{\epsilon_k\}_{k\in\mathbb{N}}$, so that

$$\left((\lambda_{(+)}^k)_j - (\lambda_{(-)}^k)_j\right) h_j(\boldsymbol{x}^k) \geq \epsilon_k \;\text{ and }\; \mu_i^k c_i(\boldsymbol{x}^k) \geq \epsilon_k, \tag{75}$$

for all $k \in \mathbb{N}$, $i \in I_p$, and $j \in I_q$. Now, due to equation (71) and inequalities (75), it is possible to conclude this part of the proof, by Proposition 5.1.

For the second part, suppose the CAKKT conditions are ensured with the primal sequence $\{\boldsymbol{x}^k\}_{k\in\mathbb{N}}$ and multipliers $\{\boldsymbol{\mu}^k\}_{k\in\mathbb{N}}$, $\{\boldsymbol{\lambda}^k\}_{k\in\mathbb{N}}$, i.e., all these sequences satisfy $\lim_{k\to\infty}(\nabla f(\boldsymbol{x}^k) + \sum_{i=1}^p \mu_i^k \nabla c_i(\boldsymbol{x}^k) + \sum_{j=1}^q \lambda_j^k \nabla h_j(\boldsymbol{x}^k)) = \boldsymbol{0}$, $\lim_{k\to\infty} \mu_i^k c_i(\boldsymbol{x}^k) = 0$ and $\lim_{k\to\infty} \lambda_j^k h_j(\boldsymbol{x}^k) = 0$.

Note that we only have to find multipliers $\{\boldsymbol{\mu}^k\}_{k\in\mathbb{N}} \subset \mathbb{R}_+^p$, $\{\boldsymbol{\lambda}_{(-)}^k\}_{k\in\mathbb{N}} \subset \mathbb{R}_+^q$, and $\{\boldsymbol{\lambda}_{(+)}^k\}_{k\in\mathbb{N}} \subset \mathbb{R}_+^q$, and a sequence $\{\boldsymbol{x}^k\}_{k\in\mathbb{N}}$ satisfying (71)–(74). For this task, we just need to identify the sequences $\{\boldsymbol{\lambda}_{(-)}^k\}_{k\in\mathbb{N}}$ and $\{\boldsymbol{\lambda}_{(+)}^k\}_{k\in\mathbb{N}}$, since $\{\boldsymbol{\mu}^k\}_{k\in\mathbb{N}}$ and $\{\boldsymbol{x}^k\}_{k\in\mathbb{N}}$ are already defined. Let $\boldsymbol{\lambda}_{(-)}^k := -(\boldsymbol{\lambda}^k)_-$ and $\boldsymbol{\lambda}_{(+)}^k := (\boldsymbol{\lambda}^k)_+$, for all $k \in \mathbb{N}$. Since, for all $i \in I_p$, $j \in I_q$ and $k \in \mathbb{N}$, we have $\boldsymbol{\lambda}^k = \boldsymbol{\lambda}_{(+)}^k - \boldsymbol{\lambda}_{(-)}^k$,

$$\begin{aligned}
|(\lambda_{(+)}^k)_j \min\{h_j(\boldsymbol{x}^k), 0\}| &\leq |\lambda_j^k h_j(\boldsymbol{x}^k)|,\\
|(\lambda_{(-)}^k)_j \min\{-h_j(\boldsymbol{x}^k), 0\}| &\leq |\lambda_j^k h_j(\boldsymbol{x}^k)|,\\
\text{and }\; |\mu_i^k \min\{c_i(\boldsymbol{x}^k), 0\} &\leq |\mu_i^k c_i(\boldsymbol{x}^k)|,
\end{aligned}$$

by the fact that the negative and positive parts of numbers are not expansive functions, the result follows. ☐

# 6 Final remarks

Inspired by the CAKKT optimality condition, this work introduced the *weighted complementary approximate Karush-Kuhn-Tucker* (WCAKKT) conditions. The new conditions have been successfully applied to the analysis of inexact restoration methods, inverse and log barrier methods, and the algorithm PACNO within the smooth context. Additionally, thanks to our main result stated by Theorem 3.1, and the WCAKKT condition, all these methods generate CAKKT sequences, and CAKKT multipliers may be obtained.

The WCAKKT conditions allow the improvement of the convergence theory of algorithms compatible with the AKKT conditions. In the analytical context, WCAKKT is proved to be as strong as CAKKT. Hence, being WCAKKT a general-purpose sequential optimality condition, and under a unified analysis, this study has provided novel ingredients that may help to understand the ability of methods in guaranteeing CAKKT sequences.

**Acknowledgements** This work was partially supported by *Conselho Nacional de Desenvolvimento Científico e Tecnológico* (CNPq) grant 305010/2020-4, and *Fundação de Amparo à Pesquisa do Estado de São Paulo* (FAPESP) grants 2013/07375-0, 2016/22989-2, 2018/24293-0, 2019/18859-4. The authors are thankful to André Belotto da Silva for the fruitful discussions regarding desingularization and monomialization, as well as to the editor and the reviewers, for the constructive comments that improved the presentation of our work.

# References

[1] Andreani, R., Fazzio, N.S., Schuverdt, M.L., Secchin, L.D.: A sequential optimality condition related to the quasi-normality constraint qualification and its algorithmic consequences. SIAM Journal on Optimization **29**(1), 743–766 (2019). URL `https://doi.org/10.1137/17M1147330`

[2] Andreani, R., Fukuda, E.H., Haeser, G., Santos, D.O., Secchin, L.D.: Optimality conditions for nonlinear second-order cone programming and symmetric cone programming. Preprint at `https://optimization-online.org/2019/10/7436/` (2019)

[3] Andreani, R., Fukuda, E.H., Haeser, G., Santos, D.O., Secchin, L.D.: On the use of Jordan algebras for improving global convergence of an augmented Lagrangian method in nonlinear semidefinite programming. Computational Optimization and Applications **79**(3), 633–648 (2021). URL `https://doi.org/10.1007/s10589-021-00281-8`

[4] Andreani, R., Haeser, G., Martínez, J.M.: On sequential optimality conditions for smooth constrained optimization. Optimization **60**(5), 627–641 (2011). URL `https://doi.org/10.1080/02331930903578700`

[5] Andreani, R., Haeser, G., Mito, L.M., Ramos, A., Secchin, L.D.: Correction to: On the best achievable quality of limit points of augmented Lagrangian schemes. Numerical Algorithms **90**, 879–880 (2022). URL `https://doi.org/10.1007/s11075-021-01241-3`

[6] Andreani, R., Haeser, G., Mito, L.M., Ramos, A., Secchin, L.D.: On the best achievable quality of limit points of augmented Lagrangian schemes. Numerical Algorithms **90**, 851–877 (2022). URL `https://doi.org/10.1007/s11075-021-01212-8`

[7] Andreani, R., Haeser, G., Schuverdt, M.L., Silva, P.J.S.: A relaxed constant positive linear dependence constraint qualification and applications. Mathematical Programming **135**(1), 255–273 (2012). URL `https://doi.org/10.1007/s10107-011-0456-0`

[8] Andreani, R., Haeser, G., Schuverdt, M.L., Silva, P.J.S.: Two new weak constraint qualifications and applications. SIAM Journal on Optimization **22**(3), 1109–1135 (2012). URL `https://doi.org/10.1137/110843939`

[9] Andreani, R., Haeser, G., Secchin, L.D., Silva, P.J.S.: New sequential optimality conditions for mathematical programs with complementarity constraints and algorithmic consequences. SIAM Journal on Optimization **29**(4), 3201–3230 (2019). URL `https://doi.org/10.1137/18M121040X`

[10] Andreani, R., Martínez, J.M., Ramos, A., Silva, P.J.S.: A cone-continuity constraint qualification and algorithmic consequences. SIAM Journal on Optimization **26**(1), 96–110 (2016). URL `https://doi.org/10.1137/15M1008488`

[11] Andreani, R., Martínez, J.M., Ramos, A., Silva, P.J.S.: Strict constraint qualifications and sequential optimality conditions for constrained optimization. Mathematics of Operations Research **43**(3), 693–717 (2018). URL `https://doi.org/10.1287/moor.2017.0879`

[12] Andreani, R., Martínez, J.M., Santos, L.T., Svaiter, B.F.: On the behaviour of constrained optimization methods when Lagrange multipliers do not exist. Optimization Methods and Software **29**(3), 646–657 (2014). URL `https://doi.org/10.1080/10556788.2013.841692`

[13] Andreani, R., Martínez, J.M., Svaiter, B.F.: A new sequential optimality condition for constrained optimization and algorithmic consequences. SIAM Journal on Optimization **20**(6), 3533–3554 (2010). URL `https://doi.org/10.1137/090777189`

[14] Andreani, R., Ramos, A., Ribeiro, A.A., Secchin, L.D., Velazco, A.R.: On the convergence of augmented Lagrangian strategies for nonlinear programming. IMA Journal of Numerical Analysis **42**(2), 1735–1765 (2022). URL `https://doi.org/10.1093/imanum/drab021`

[15] Andreani, R., Ramos, A., Secchin, L.D.: Improving the global convergence of inexact restoration methods for constrained optimization problems (2022). URL `http://www.optimization-online.org/DB_HTML/2022/03/8851.html`

[16] Belotto da Silva, A., Bierstone, E.: Monomialization of a quasianalytic morphism. Preprint at `https://arxiv.org/pdf/1907.09502.pdf` (2019)

[17] Bertsekas, D.P.: Nonlinear Programming, 2 edn. Athena scientific optimization and computation series. Athena Scientific, Belmont, Massachussets (1999)

[18] Bierstone, E., Milman, P.D.: Semianalytic and subanalytic sets. Publications Mathématiques de l'Institut des Hautes Études Scientifiques **67**(1), 5–42 (1988). URL `https://doi.org/10.1007/BF02699126`

[19] Birgin, E.G., Martínez, J.M.: Practical Augmented Lagrangian Methods for Constrained Optimization. Society for Industrial and Applied Mathematics, Philadelphia, PA (2014). URL `https://doi.org/10.1137/1.9781611973365`

[20] Cossart, V., Jannsen, U., Saito, S.: Desingularization: Invariants and Strategy, *Lecture Notes in Mathematics*, vol. 2270, 1 edn. Springer Cham (2020). DOI 10.1007/978-3-030-52640-5

[21] Cutkosky, S.D.: Monomialization of Morphisms from 3-Folds to Surfaces, *Lecture Notes in Mathematics*, vol. 1786, 1 edn. Springer-Verlag Berlin Heidelberg (2002). DOI 10.1007/b83848

[22] Dempe, S., Kalashnikov, V., Pérez-Valdés, G.A., Kalashnykova, N.: Bilevel Programming Problems: Theory, Algorithms and Applications to Energy Networks, 1 edn. Energy Systems. Springer-Verlag Berlin Heidelberg (2015). DOI 10.1007/978-3-662-45827-3

[23] Dempe, S., Zemkoho, A.: Bilevel Optimization: Advances and Next Challenges, *Springer Optimization and Its Applications*, vol. 161, 1 edn. Springer Nature Switzerland AG (2020). DOI 10.1007/978-3-030-52119-6

[24] Feehan, P.M.N.: Resolution of singularities and geometric proofs of the Łojasiewicz inequalities. Geometry & Topology **23**(7), 3273–3313 (2019). URL `https://doi.org/10.2140/gt.2019.23.3273`

[25] Forsgren, A., Gill, P.E., Wright, M.H.: Interior methods for nonlinear optimization. SIAM Review **44**(4), 525–597 (2002). URL `https://doi.org/10.1137/S0036144502414942`

[26] Gill, P.E., Kungurtsev, V., Robinson, D.P.: A stabilized SQP method: global convergence. IMA Journal of Numerical Analysis **37**(1), 407–443 (2016). URL `https://doi.org/10.1093/imanum/drw004`

[27] Gill, P.E., Kungurtsev, V., Robinson, D.P.: A shifted primal-dual penalty-barrier method for nonlinear optimization. SIAM Journal on Optimization **30**(2), 1067–1093 (2020). URL `https://doi.org/10.1137/19M1247425`

[28] Gill, P.E., Robinson, D.P.: A globally convergent stabilized SQP method. SIAM Journal on Optimization **23**(4), 1983–2010 (2013). URL `https://doi.org/10.1137/120882913`

[29] Gomes-Ruggiero, M.A., Martínez, J.M., Santos, S.A.: Spectral projected gradient method with inexact restoration for minimization with nonconvex constraints. SIAM Journal on Scientific Computing **31**(3), 1628–1652 (2009). URL `https://doi.org/10.1137/070707828`

[30] Haeser, G.: On the global convergence of interior-point nonlinear programming algorithms. Computational & Applied Mathematics **29**(2), 125–138 (2010). URL `https://doi.org/10.1590/S1807-03022010000200003`

[31] Haeser, G.: A second-order optimality condition with first- and second-order complementarity associated with global convergence of algorithms. Computational Optimization and Applications **70**(2), 615–639 (2018). URL `https://doi.org/10.1007/s10589-018-0005-3`

[32] Helou, E.S., Santos, S.A., Simões, L.E.A.: Analysis of a new sequential optimality condition applied to mathematical programs with equilibrium constraints. Journal of Optimization Theory and Applications **185**(2), 433–447 (2020). URL https://doi.org/10.1007/s10957-020-01658-1

[33] Helou, E.S., Santos, S.A., Simões, L.E.A.: A new sequential optimality condition for constrained nonsmooth optimization. SIAM Journal on Optimization **30**(2), 1610–1637 (2020). URL https://doi.org/10.1137/18M1228608

[34] Kanzow, C., Raharja, A.B., Schwartz, A.: Sequential optimality conditions for cardinality-constrained optimization problems with applications. Computational Optimization and Applications **80**(1), 185–211 (2021). URL https://doi.org/10.1007/s10589-021-00298-z

[35] Kashiwara, M., Schapira, P.: Sheaves on Manifolds, *Grundlehren der mathematischen Wissenschaften*, vol. 292, corrected second printing edn. Springer-Verlag (1994). DOI 10.1007/978-3-662-02661-8

[36] Kovacevic, R.M., Pflug, G.C., Vespucci, M.T.: Handbook of Risk Management in Energy Production and Trading, *International Series in Operations Research & Management Science*, vol. 199, 1 edn. Springer New York, NY (2013). DOI https://doi.org/10.1007/978-1-4614-9035-7

[37] Krulikovski, E., Ribeiro, A., Sachine, M.: On the weak stationarity conditions for mathematical programs with cardinality constraints: a unified approach. Applied Mathematics & Optimization **84**, 3451–3473 (2021). URL https://doi.org/10.1007/s00245-021-09752-0

[38] Machado, F.C., de Oliveira Filho, F.M.: Improving the semidefinite programming bound for the kissing number by exploiting polynomial symmetry. Experimental Mathematics **27**(3), 362–369 (2018). URL https://doi.org/10.1080/10586458.2017.1286273

[39] Martínez, J.M., Pilotta, E.A.: Inexact-restoration algorithm for constrained optimization. Journal of Optimization Theory and Applications **104**(1), 135–163 (2000). URL https://doi.org/10.1023/A:1004632923654

[40] Martínez, J.M., Svaiter, B.F.: A practical optimality condition without constraint qualifications for nonlinear programming. Journal of Optimization Theory and Applications **118**(1), 117–133 (2003). URL https://doi.org/10.1023/A:1024791525441

[41] Maurya, J.K., Mishra, S.K.: Strong complementary approximate Karush-Kuhn-Tucker conditions for multiobjective optimization problems. Yugoslav Journal of Operations Research **32**(2), 219–234 (2022). URL https://doi.org/10.2298/YJOR210315024M

[42] Nocedal, J., Wright, S.J.: Numerical Optimization, 2 edn. Springer, New York, NY, USA (2006)

[43] Prado, R.W., Santos, S.A., Simões, L.E.A.: On the convergence analysis of a penalty algorithm for nonsmooth optimization and its performance for solving hard-sphere problems. Numerical Algorithms (2022). URL https://doi.org/10.1007/s11075-022-01287-x

[44] Sussmann, H.J.: Real-analytic desingularization and subanalytic sets: An elementary approach. Transactions of the American Mathematical Society **317**(2), 417–461 (1990). URL http://www.jstor.org/stable/2001472